. 475
959

Psychological Monographs

General and Applied

The Ontogeny of Certain Logical Operations: Piaget's Formulation Examined by Nonverbal Methods

By

Martin D. S. Braine

New York University

Braine

Price $1.50

l. 73
o. 5

Edited by Norman L. Munn

Published by the American Psychological Association, Inc.

Psychological Monographs:

General and Applied

Combining the *Applied Psychology Monographs* and the *Archives of Psychology* with the *Psychological Monographs*

NORMAN L. MUNN, Editor

Department of Psychology, Bowdoin College
Brunswick, Maine

Consulting Editors

ANNE ANASTASI
FRANK A. BEACH
ARNOLD M. BINDER
W. J. BROGDEN
ROBERT R. BUSH
JOHN F. DASHIELL
JAMES J. GIBSON
D. O. HEBB
EDNA HEIDBREDER
FRANCIS W. IRWIN
JAMES J. JENKINS
HAROLD E. JONES
DANIEL KATZ
BOYD MCCANDLESS
DONALD W. MACKINNON
QUINN MCNEMAR
LORRIN A. RIGGS
CARL R. ROGERS
RICHARD L. SOLOMON
ROSS STAGNER

Manuscripts and correspondence on editorial matters should be sent to the Editor. *Psychological Monographs* publishes comprehensive experimental investigations and programmatic studies which do not lend themselves to adequate presentation as journal articles. Major space is given to the author's original contribution; introductory and bibliographic materials, as well as statistical tables and graphs, must be kept within reasonable bounds. Tables, graphs, and appendix materials which deal with detail not essential to adequate presentation of the findings may be made available through the American Documentation Institute—for details of this procedure, see the APA *Publication Manual*. Preparation of manuscripts for publication as monographs should follow the procedure given in the APA *Publication Manual*. Publication in *Psychological Monographs* is free of cost to the author, except in cases where early publication is requested or author's alterations are made in galley proofs.

ARTHUR C. HOFFMAN, Managing Ed.; HELEN ORR, Promotion Mgr.; SADIE J. DOYLE, Editorial Asst.

Correspondence on business matters should be addressed to the American Psychological Association, Inc., 1333 Sixteenth St., N.W., Washington 6, D.C. Address changes must arrive by the 10th of the month to take effect the following month. Undelivered copies resulting from address changes will not be replaced; subscribers should notify the post office that they will guarantee third-class forwarding postage.

COPYRIGHT, 1959, BY THE AMERICAN PSYCHOLOGICAL ASSOCIATION, INC.

Vol. 73, No. 5 Whole No. 475, 1959

Psychological Monographs: General and Applied

THE ONTOGENY OF CERTAIN LOGICAL OPERATIONS:

PIAGET'S FORMULATION EXAMINED BY NONVERBAL METHODS[1]

MARTIN D. S. BRAINE[2]

New York University

THE EXPERIMENTS reported are concerned with the work of Piaget on the development of intelligence in children. Piaget's theory has two aspects. First, it is a theory of intelligence: Piaget believes that the development of intelligence consists in the development of an ability to perform logical operations. Second, Piaget makes specific statements about the ages at which certain types of reasoning develop. He claims that as children grow older certain specified groups of logical operations develop in the average child at given ages. In his studies the principal ages of transition are reported to be at 7 and 11 yr., approximately. At around age 7, the operations of the spatial and class interpretations of the Boolean calculus make their appearance in children's thinking; at around age 11, the operations of the propositional interpretation appear.[3] The appearance of the operations of the spatial interpretation are revealed in Piaget's studies by changes at around age 7 yr. in the child's manner of using measuring instruments and in his performance on a number of tasks supposedly demanding conception of, or reasoning about, spatial relations (Piaget & Inhelder, 1956; Piaget, Inhelder, & Szeminska, 1948). The operations of the class interpretation are revealed by the development of number concepts and by the disappearance of certain anomalies in the child's class concepts (Piaget & Szeminska, 1952). The appearance of propositional operations at age 11 is attested, according to Piaget, by the child's performance on a variety of reasoning tasks (Piaget, 1955; Piaget & Inhelder, 1953).

The notion that the nature of intelligence lies in the performance of logical operations is not new in psychology. It is clearly implicit in Spearman's (1923) identification of intelligence with "eduction of relations"

[1] This study is based upon a doctoral dissertation submitted to the Department of Psychology, New York University Graduate School of Arts and Science. The writer wishes to express his thanks to Elsa Robinson who sponsored the research, and to H. L. Teuber for his invaluable suggestions and criticisms. The writer is also indebted to Marjorie Cornell and Elsie Stange for making available the subjects and the facilities of their nursery schools. The work was partially supported by the Commonwealth Fund of New York through their grant to H. L. Teuber and the Psychophysiological Laboratory, New York University-Bellevue Medical Center.

[2] Now at State University of New York.

[3] For the calculus and its various interpretations, see, for example, Langer (1953). Piaget's logical system differs from the Boolean calculus, but the differences can be neglected for the purposes of this study. The interested reader may refer to his treatise on logic (Piaget, 1949).

The operations of the Boolean calculus are called *conjunction* and *disjunction*. A conjunction of classes is denoted in the English language by the qualification of a noun by an adjective, as, for instance, "white sheep" (i.e., objects that are *both* white, *and also* sheep). A disjunction of classes is often expressed by the words *and* and *or*, as, for instance, "men *and* women," "barren mountains, *or* fertile valleys."

A spatial conjunction is expressed in the two following examples: "The United Nations building is located at 42nd Street *and* the East River Drive," "a straight line is the region *common to* two intersecting planes." Spatial disjunction indicates partition or composition, e.g., "Canada *and* the United States make up North America."

A conjunction of propositions asserts the joint truth of two propositions, often by the use of the word *and*, for example, "Fabius was born at the rising of the dog-star, *and* Fabius will perish in the sea." A disjunction of propositions carries the meaning "either . . . or . . . (or both)."

and "eduction of correlates."[4] Moore (1929) seems to have had a similar view. Attempts to find logical models for schizophrenic thinking (Arieti, 1948; Von Domarus, 1944) also implicitly make the assumption that the only proper model for "normal" reasoning was described by Aristotle.

Piaget's theory differs from those of previous investigators, first, in that it aims to give a description of the *development* of intelligence, and second, in that his logical model for reasoning is worked out in great detail and is derived from modern symbolic logic (Piaget, 1949). Piaget attempts to support his theory with a good deal of evidence, although the validity of some of it has been questioned (Fraisse & Vautrey: 1952a, 1952b).

Testing Piaget's theory in detail will obviously prove a monumental task, not only because of its tremendous scope, but also because some of his concepts do not have a sufficiently firm basis in experimental operations. The investigation reported here is intended as an attempt to examine the validity of Piaget's theory within a limited context. His views on the development of length measurement and of concepts of order in children have been singled out for study. The development of length measurement is discussed in the next section and order concepts in the section following. In both these sections special attention is paid to Piaget's views on the specific ages at which particular types of responses develop.

The ontogenesis of measurement provides a particularly convenient point of departure since the logic of measurement has been well analyzed by workers interested in the bases of science (Bergmann & Spence, 1944; Campbell, 1920; Nagel, 1930). In the next section an experiment is reported in which the development of a particular inferential response in children is traced. This response is the inference that a stick, A, is longer than another stick, C, when it has been demonstrated that A is longer than a measuring rod, B, and that B is in turn longer than C. In logical symbolism this may be written $A > B.B > C. \supset .A > C$.[5] According to Nagel and other logicians this inference is one of those fundamental in length measurement (i.e., all length measurement tacitly assumes the validity of this inference).

In the section entitled The Development of a Concept of Position Order, the development of discrimination of position order in children is studied. An experiment is reported in which *S* is invited to match a sample card on which colored discs have been pasted in a given order, say, red-gray-green, with one of three other cards in which the order of the colors is gray-red-green, red-green-gray, and red-gray-green.

In the section which follows, the relationship between the development of the inferential response $A > B.B > C. \supset .A > C$ and of the discrimination of position order is investigated. Piaget believes that the development of inferences such as $A > B.B > C. \supset .A > C$ is one aspect of the development of a more general ability to perform logical operations. The task Piaget has set himself is to give a detailed description of these logical operations and of their formal interrelations. He believes that this will enable us "to construct by means of the algebra of logic a deductive theory to explain some of the experimental findings of psychology" (Piaget, 1953b, p. 26). Although Piaget does not make this explicit, a basic postulate of his theory of intelligence certainly appears to be that operations, inferences, or concepts which are logically interdependent (i.e., the definition of the one involves reference to the other, or both derive from the same interpretation of a logical calculus), will be found to develop in association with each other in children's thinking. It may be noted that experimental refutation of Piaget's views about the specific ages at which particular groups of operations develop does not refute this postulate. With this postulate it

[4] Piaget (1953b) presents a derivation of these concepts from his own theory.

[5] To be read: "*If* (A is longer than B) *and* (B is longer than C) *then* (A is longer than C).

may be predicted that order concepts and the inference $A > B.B > C. \supset .A > C$ develop in association in children's thinking.

If the postulate that logically interdependent operations develop in association is generally valid, it follows that there should be periods of transition in the development of intelligence, during which particular groups of operations develop. If such ages of transition exist—and they may not be at 7 and 11 yr. as Piaget claims—then changes in the nature of intelligence during childhood should be observable and reflected in a changing factorial composition of intelligence tests. Failure to find such changes would cast serious doubt on any theory which seeks to relate intellectual development to the performance of logical operations. In the section on ages of transition, this study concludes with an examination of the evidence for such changes reported in the literature.

The Development of Length Measurement in Children

The part played by the logic of measurement in Piaget's description of length-measuring behavior in children is considered. Piaget's experiments are discussed and certain criticisms of his techniques are suggested. A study which uses different techniques to study the development of certain of the inferential processes involved in measurement is then reported. The results are compared with Piaget's.

Theoretical Considerations

According to many logicians, measurement is possible in a dimension only when relations between objects in the dimension fulfill certain axioms, the so-called axioms of measurement (Bergmann & Spence, 1944; Nagel, 1930). From these, certain laws may be derived. For example, the truth of the following seven laws is always tacitly assumed whenever length is measured.[6]

$$A = B.B = C. \supset .A = C$$
$$A = B.B > C. \supset .A > C$$
$$A = B.B < C. \supset .A < C$$
$$A > B.B = C. \supset .A > C$$
$$A > B.B > C. \supset .A > C$$
$$A < B.B = C. \supset .A < C$$
$$A < B.B < C. \supset .A < C$$

When length is measured, it is also assumed that lengths are additive, i.e., any length is composed of shorter lengths added together. Consequently, some arbitrary unit (e.g., in., cm.) may be defined and the length of a measured object is the number of units which have to be added together to compose another object equal in length to the first. It is the existence of this additive operation that, among other things, enables one to talk meaningfully of the length of a path which is not straight, e.g., the perimeter of a triangle whose length is the sum of the lengths of the sides.

Piaget's description of the development of measurement in children consists of an account of the development of children's understanding and use of these laws.

Piaget treats these laws not only as empirical statements which are demonstrably valid, but as behavioral processes initiated by the child *S*. For example, the expression (A + B) can be regarded either as a length, or as a record of the behavior of some *S*: "*S* picks up B and places one end of B next to one end of A." If *S* does not displace B physically, but does so only implicitly or "in his imagination," then the expression (A + B) can be regarded as representing an implicit displacement or operation performed by *S* on A and B.

In most studies of length judgment in children, the *S*s have been asked to compare two lengths. The method used by *S* to arrive at a judgment is likely to depend both on the stimulus conditions and on the manner in which *S* is obliged to communicate his judgment to *E*.

The importance of the stimulus conditions in length judgment may be illustrated by two prototypic examples. In the first example, a length A in one room is to be compared with another somewhat shorter length C in another room by the following procedure: in the first room a convenient third object, B, is juxtaposed with A and it is found that A is longer than B; B is then carried into the other room and juxtaposed with C and it is found that B is longer than C; it is concluded that A is longer than C. For the second example, suppose that the stimulus conditions are altered so that A and C are in the same room standing upright next to each other on the top of a table; it is then immediately clear that A extends upwards beyond the highest point of C, i.e., that A is longer than C.

Let us say that in the first case A is "inferred" to be longer than C, and in the second case that A is "perceived" to be longer than C. This appears to be not inconsistent with conventional usage of these words in the English language. In the remainder of this section the word *inference* is used to refer to the kind of method by which the length judgment is arrived at in the first example, and the judgment itself is called an *inferential response*.

[6] A, B, and C represent any three lengths and ">," "<," and "=" designate the relations "longer than," "shorter than," and "equal in length to," respectively

The terms *perception* and *discrimination* will often be used to designate the length judgment in the second example.

Critique of Piaget's Techniques

In Piaget's work the stimulus conditions are rarely identical with those of the two prototypic examples above; distinguishing between an inferential and a discriminatory response therefore becomes a diagnostic problem. In his experiments the stimuli to be compared have sometimes been solid objects and sometimes the paths traced by moving objects. The stimuli have been straight or bent to form a "zig-zag," and the spatial relations between the stimuli have been varied. The response has sometimes been a judgment expressed verbally, and sometimes *S* has been asked to construct a length equal to a given length, with or without uncalibrated measuring rods available. In all experiments either *E* or *S* has used words such as *big, little, short, long, same length, measure,* etc., so that a certain vocabulary development is presupposed.

The standardization of the Binet test (see Terman & Merrill, 1937) indicates that average children between the ages of 3–0 and 3–6 are able to obey a verbal instruction to pick out the longer of two sticks 2 in. and 2½ in. in length at least 82% of the time, when the sticks are placed parallel so that the longer overlaps the shorter at both ends. It is an open question whether the word *longer* is understood at this age to apply to other spatial relations between the sticks.

In two experiments by Piaget (Piaget, 1946, Ch. 3; Piaget et al., 1948, Ch. 6), the *S*s were aged 4–8 yr. The *E* moved a bead a certain distance along a wire and *S* was required to move another bead the same distance along a second wire.[7] The shape and relative positions of the wires together with the typical errors of the youngest *S*s are shown in Fig. 1. The errors made by most of the *S*s aged 4–6 yr. were always of the same nature: they moved their bead to a position, P′, directly opposite the position, P, of *E*'s bead, thus failing to take account of the differing points of departure of the two beads. The responses of these *S*s were thus correct only in arrangement "B" in Fig. 1. The *S*s of intermediate age were correct in arrangements "A" and "C," but not in "D" or "E." The *S*s aged 7 and 8 yr. were correct in all arrangements. Following their response all *S*s were given an unmarked ruler to measure the distances. The youngest *S*s just laid the ruler across the two beads perpendicular to the wires. Not until after age 7 yr. were the *S*s able to measure adequately. Piaget interpreted his findings to indicate that his youngest *S*s identified "length" with "final position occupied"; the typical errors indicated the "finalism" which is characteristic of children of this age, and which reflects the fact that the logical operations basic to measurement are not yet available to them.

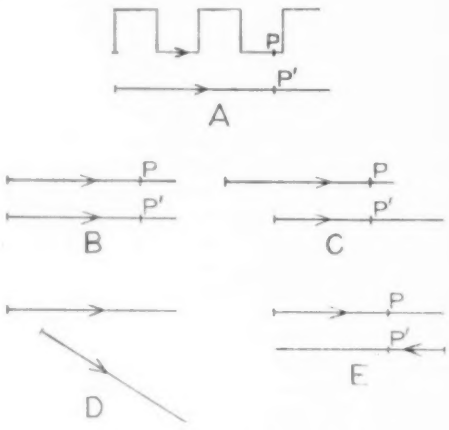

FIG. 1. Arrangements of wires used by Piaget. The arrow indicates direction of travel and P and P′ the points of arrival of the beads moved, respectively, by *E* and by the youngest group of *S*s (with whom Arrangement "D" was not used).

Fraisse and Vautrey (1952a) criticize Piaget's procedure and, in a study which also employed the paths traced by moving objects as stimuli, they found no evidence for "finalism" in young children's perception of length. Their *S*s were aged 4–6 to 5–6,

[7] "Veux-tu faire marcher l'autre tram pour qu'il fasse juste le même long chemin?" (Piaget, 1946, p. 56).

the age of Piaget's youngest *S*s. Fraisse and Vautrey found that *S*s who were told to indicate the starting positions of the objects responded correctly more often than *S*s whose response was not prefaced by this instruction. Moreover, *S*s who made errors had usually indicated the starting position incorrectly, their response being appropriate to the starting position indicated. The authors suggest that the errors typical of Piaget's youngest *S*s reflect not a peculiarity of their conception of length, but their failure to notice or to recall the starting positions of the objects.

If the errors made by Piaget's *S*s do not constitute evidence to support his diagnosis of their length judgments as "pre-logical," then the only support for this diagnosis lies in the inappropriateness of the measuring behavior of his *S*s.

The inception of measurement is examined in greater detail in another of Piaget's studies (Piaget, 1953a; Piaget et al., 1948, Ch. 2). Piaget invited children to build with blocks a tower equal in height to a tower already built by *E*. However, the table on which *E*'s tower stood was higher than the table on which *S* was to build his tower and was some distance from it. Sticks longer, shorter, and equal to the height of the model were available to *S*s. Piaget summarizes his results as follows:

> The youngest children build up the second tower to the same visual level as the first, without worrying about the difference in height of the tables. They compare the towers by stepping back and sighting them. At a slightly more advanced stage a child lays a long rod across the tops of the two towers to make sure that they are level. Somewhat later he notices that the base of his tower is not at the same level as the model's. He then wants to place his tower next to the model on the same table to compare them. Reminded that the rules of the game forbid him to move his tower, he begins to look around for a measuring standard. Interestingly enough, the first that comes to his mind is his own body. He puts one hand on top of his tower and the other at its base, and then, trying to keep his hands the same distance apart, he moves over to the other tower to compare it. Children of about the age of six often carry out this work in a most assured manner, as if their hands could not change position on the way! Soon they discover that the method is not reliable, and then they resort to reference points on the body. The child will line up his shoulder with the top of his tower, mark the spot opposite the base on his thigh with his hand and walk over to the model to see whether the distance is the same.
>
> Eventually the idea of an independent measuring tool occurs to the child. His first attempt in this direction is likely to be the building of a third tower next to and the same height as the one he has already erected. Having built it, he moves it over to the first table and matches it against the model; this is allowed by the rules. The child's arrival at this stage presupposes a process of logical reasoning. If we call the model tower A, the second tower C, and the movable tower B, the child has reasoned that B = C and B = A, therefore A = C.
>
> Later the child replaces the third tower with a rod, but at first the rod must be just the same length as the height of the tower to be measured. He then conceives the idea of using a longer rod and marking the tower height on it with his finger. Finally, and this is the beginning of true measurement, he realizes that he can use a shorter rod and measure the height of the tower by applying the rod a certain number of times up the side.
>
> The last discovery involves two new operations of logic. The first is the process of division which permits the child to conceive that the whole is composed of a number of parts added together. The second . . . enables him to apply one part upon others and thus to build a system of units (Piaget, 1953a, pp. 77–78).

It can be seen that Piaget's diagnosis of a child's ability to perform the inferences basic to measurement is based entirely on his measuring behavior. However, it seems to the writer that a child's measuring behavior may reflect his knowledge of measuring technique as much as it does his ability to reason logically about length. Piaget's 5- and 6-yr.-old *S*s tended to use their own bodies as measuring rods. The human body is, of course, a clumsy and inadequate measuring instrument, yet the fact that a young child does not immediately perceive its imperfections and select a better tool may not necessarily indicate a deficiency in his logic. One may recall that in the cultural history of measurement it required a considerable period of time before adult human beings came to adopt more objective units of measurement than parts of the human body. Consider the behavior of the child who "measures" the height of a tower by placing one hand at its base and the other at its summit. Piaget claims that such behavior indicates deficient understanding of the logic of measurement, because the child

fails to realize that it is impossible to keep the distance between the hands constant as he walks from one tower to the other: the child's measuring rod is elastic. However, the opposite point of view can be argued. To the writer such behavior suggests that the child has a rather sound knowledge of the logic of measurement, but is woefully inexpert in its practice! Such behavior appears to show a good knowledge of the seven paradigmatic rules of inference involved in measurement enumerated earlier.

Another factor which may be important in Piaget's studies of measurement is the child's motivation. Piaget relies on the inherent interest of the tasks for the child. However, it is possible that the older children are more interested than the younger *S*s in measurement and the additional accuracy it makes possible.

In another experiment (Piaget et al., 1948, Ch. 5) designed to study children's understanding of the additive operation involved in length measurement, the stimuli were two "roads" composed of match sticks of unequal length placed end to end. Two sets of sticks were initially lined up parallel to each other, and without overlap, to form two equally long lines. The *S* was asked whether the "roads" were the same length or not. Even the youngest *S*s were able to respond "Yes" to this question.

One line of sticks was then altered so that the "road" was no longer straight, but formed a "zig-zag." The *S* observed the alteration. He was then asked again whether the "roads" were the same length or not. This question was sometimes rephrased, the child being asked, if an ant were to walk along each road, whether one ant would walk farther than the other.

It was found that the youngest children (aged 4–6 yr.) did not consider the lines the same length after the rearrangement had taken place; they considered either the straight line to be longer because it extended beyond the other one, or else they thought the "zig-zag" line to be longer because of the turns. There was an intermediate stage where some rearrangements were considered to alter the length and some were not. Around ages 7 and 8 yr. (Stage III), the line was known to be the same length irrespective of the rearrangement of its parts.

This experiment is one of the rare ones in which Piaget reports some results numerically. Fifty-nine children between the ages of 4 and 9 were examined. Responses classified in Stage III were shown by only 10% of those aged 6–0 to 7–0, by 50% of those between 7–0 and 7–6, and by 75% of those between 7–6 and 8–6.

No manipulation of measuring instruments is required of the child in this experiment, but Piaget's questions are very complex for young children. Indeed, the use of words such as *long* or *length* (which appear in the questions) to refer to lines which are not straight presupposes comprehension of the additive operation involved in measurement, since the length of a line which is not straight is defined as the sum of the lengths of its straight parts. Comprehension of Piaget's questions by the child would therefore of itself indicate understanding of the additive character of lengths. Piaget's results can thus only be regarded as equivocal, since it is unclear precisely what he is studying. The development of this additive operation could perhaps be examined more adequately if the experiment were redesigned as a length-discrimination task in which selection of the longer of two sticks (at least one of which is bent to form a zig-zag) led to, say, a candy reward. At whatever age the additive operation develops (if there is such an age) one would expect an increase in skill in visualizing a line as composed of displaceable parts, and therefore a sudden improvement in the accuracy of discrimination at this age.

Vocabulary development may well be a factor in many of Piaget's experiments. Its confounding effects cannot be eliminated unless nonverbal methods of presenting the tasks to the *S*s are used. The notion of employing nonverbal techniques has methodological importance. Historically, one of the important contributions of comparative psychology has been the consistency with which processes of reasoning or concept formation, which at one time were considered by some to be peculiar to the human species, have been elicited in lower animals, once the ap-

propriate nonverbal tasks have been designed. Field's (1932) study of concept formation in the rat provides an example of this. The success of Harlow (1951), and Evarts and Nissen (1953) in teaching subhuman primates to solve conditional matching problems employing the Weigl principle is a more recent instance, which suggests that the kind of thinking called "abstract" by Goldstein and Scheerer (1941) can occur in primates other than man.[8]

No theory which postulates levels of conceptual development can be regarded as definitively established when the supporting data are obtained through extensive verbal communication with *S*s who differ in their ability to verbalize. This principle applies not only when the human species is compared with phylogenetically lower animals, but also, within the human species, when aphasics—and perhaps also schizophrenics—are compared with "normals," and in ontogenetic studies of young children.

The main purpose of using nonverbal techniques is therefore not to discredit the notion of "levels" or "stages" of development, but to provide more convincing evidence for them.[9]

Conservatively interpreted, the evidence suggests the following conclusions:

1. In the absence of measurement, the accuracy of length judgment in young children varies with the nature and the spatial arrangement of the stimuli judged. Where the lengths are the paths traced by moving objects, failure to notice or to recall the relative starting positions is an important source of error.

2. Selection and manipulation of instruments for measuring length becomes increasingly adequate between the ages of 4 and 8 yr. in the average child who has had no special tuition.

3. There is reason to question Piaget's belief that performance of the inferences and logical operations basic to measurement emerges at the approximate age of 7 yr. in the average child. Piaget's findings may reflect his methods of investigation.

To evaluate this criticism of Piaget, a length-discrimination task was designed in which the difference in length between the stimuli could not readily be perceived but had to be inferred, using the rule $A > B.B > C. \supset .A > C$. This rule has the status of an axiom in the logic of measurement. The *S* did not manipulate any measuring instruments during the task so that skill in the use of measuring rods was not a controlling factor. Every time the child made a correct response he obtained a small piece of candy, so that he was not motivated only by the inherent interest of the task.

Tracing the development of this inferential response in children makes it possible to subject to empirical test Piaget's view that the inferences fundamental to length-measurement develop in children at around age 7 yr.

Method

Subjects

The *S*s were children from two day nurseries in New York City, partially supported by the Division of Day Care of the New York City Department of Welfare. There were 41 *S*s, 18 boys and 23 girls, who ranged in age from 3–6 to 7–0. The distribution by age is shown in Table 1. The children in these nurseries are almost all children of working mothers who live in low-income housing projects in New York's lower East Side. Although the socioeconomic status and education of the parents are below average for New York City, in the opinion of the nursery authorities and of the writer, it is doubtful that the children are below the country-wide average in general intelligence. IQs are available for very few of the children, but those that are range from 89 to 118 and thus are in harmony with this general impression.

Of the 41 *S*s, 37 also received the order-discrimination task described in the next section. To 16 of these, the order-discrimination task was administered before the task now described.

[8] Moreover, deficits in "abstraction" following brain injury may not be the peculiar privilege of man! Cf. Riopelle, Alper, Strong, and Ades (1953).

[9] The writer is indebted to H. L. Teuber for suggesting the use of nonverbal techniques. In numerous conversations he has persistently reiterated that, if a theory postulates an alteration of thought (rather than language) processes either during development or as a result of brain injury or disease, then the possibility of demonstrating this alteration on nonverbal tests is clearly implied.

Apparatus

The apparatus consisted of 15 upright pieces of wood painted a turquoise color, screwed to a flat base painted black. Each of these could be placed to rest on a low orange pedestal, constructed so that a piece of candy could be concealed in the pedestal beneath the base of the upright. The apparatus also contained two orange measuring sticks. The 15 uprights are shown in Fig. 2. Two of these 15 are shown in Fig. 3 being measured in turn with a measuring stick. The longer of the two measuring sticks, which measured 11⅝ in. and 11¾ in., is shown in Fig. 3. Each of the measuring sticks was designed with a spigot on top which passed over the shorter of the two uprights, thus clearly revealing the relative lengths of the uprights and measuring stick. This spigot can be seen best in the left photograph of Fig. 3.

The purpose of constructing the uprights with arms at their upper and lower extremities was to increase the difficulty of making accurate discriminations of their relative lengths by direct visual comparison. It was expected that the arms would induce a slight degree of illusion (Müller-Lyer effect) in the apparent length of the uprights.

Procedure

The task consisted of a series of discrimination problems with the stimuli (the uprights) presented in pairs. There were three phases. In the first phase the uprights used were always clearly different in height. Half the children learned to obtain a reward by selecting the longer upright, the remainder by selecting the shorter upright. In the second phase *S* was required to find the rewarded upright when the two uprights in each trial were

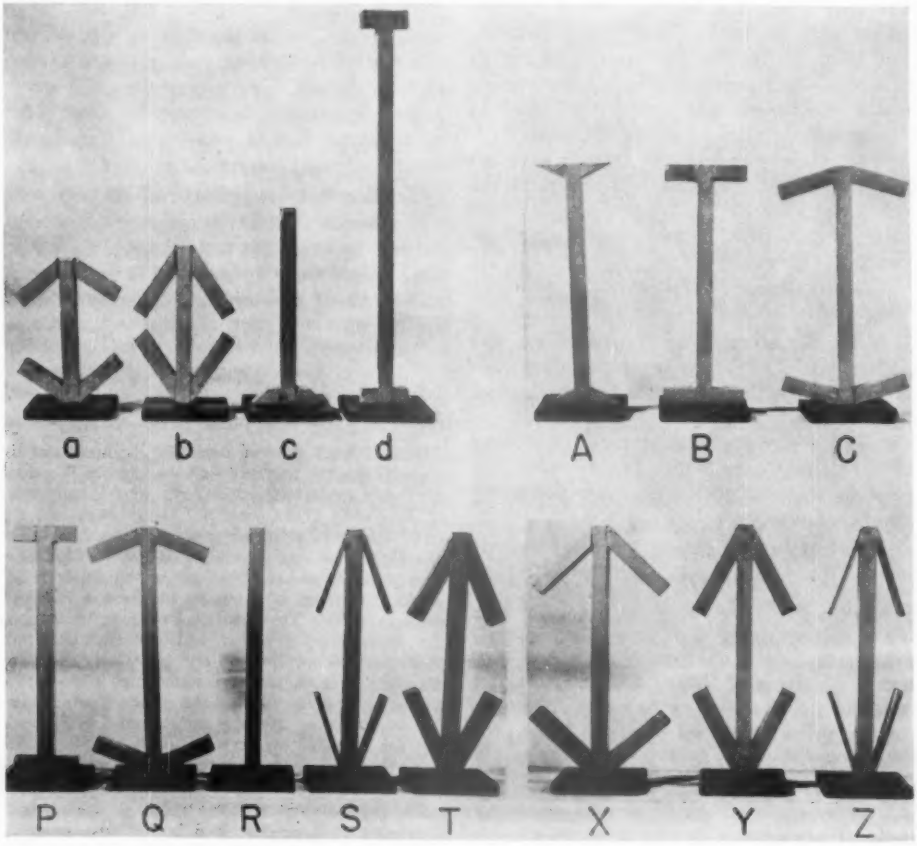

FIG. 2. Stimuli used in the measurement task. a, b, c, and d were used only in training and retention trials. A, B, and C were 11⅞ in. tall, P, Q, R, S, and T 11⅞ in., and X, Y, and Z 11⅝ in.

objectively, but not discriminably, different in length, and where each was compared in turn, by *E*, with a measuring stick of intermediate size. The *S* had to *infer* which was the longer (or shorter) upright, by the rule A > B.B > C. ⊃ .A > C. In the third phase the question, "Which is the taller [or shorter]?" was posed verbally as *E* measured the uprights with the measuring stick. Thus, in Phase 1 each child was taught the relation "longer than" (or "shorter than"). Phases 2 and 3 tested whether this relation was transitive or intransitive for the child: in Phase 2 a nonverbal method was used, and in Phase 3 the nonverbal method was supplemented by verbal instructions. The precise details of procedure in the three phases were as follows.

Phase 1: Training. The *S* and *E* sat opposite each other, with the long edges of a low rectangular table between them. The *S* was first shown how to lift one of the uprights and find candy under it. The uprights were then presented in pairs. Half the children in each age group always found candy under the taller upright, the remainder under the shorter upright. The uprights in each pair were clearly different in height: during Phase 1 no two of the uprights designated in Fig. 2 by capital letters were ever paired together. The uprights designated d and P in Fig. 2 constituted the first pair of training stimuli. This pair was presented in successive trials until *S* had achieved eight successive correct comparisons. The pair d-P was then removed and a-X substituted for it. When *S* had performed six successive correct comparisons of a and X, these were removed and d-P reinstated. If *S* did not remember whether the candy was under d or P, he was retrained. Pairs d-P and a-X were then presented on successive trials until both discriminations were retained simultaneously. The pairs a-P, P-d, d-X, and X-a were then presented in successive trials until all four discriminations had been learned and retained simultaneously. If *S* proved unusually slow in learning to transpose the concept taught, these pairs were presented individually in successive trials until learned. Then c-R was introduced and *S* was trained on this if necessary. Then the pairs b-c, c-R, R-d, d-Y, Y-c, c-a were presented successively until all the discriminations involved were retained simultaneously. When this point was reached, uprights designated by lower-case letters in Fig. 2 were paired at random in successive trials with uprights designated by capital letters until eight successive correct comparisons of different pairs of uprights were made. The *S* was then considered to have learned whichever of the concepts "longer than" or "shorter than" he had been taught and Phase 2 was initiated. When the training phase required more than one session with an *S*, pairs of stimuli were presented in subsequent sessions in the same sequence as they had been presented in the first sessions: that is, the child was always trained from the beginning again, although the success criteria for the discriminations learned in prior sessions were somewhat relaxed.[10]

During Phase 1, the position of the correct upright was varied, but it was not always varied randomly. It was found that random variation of the correct position was an ineffective training technique since it generally provided accidental partial reinforcement of an *S*'s wrong hypotheses or habits. Many *S*s tended to reach for the upright nearest their preferred hand. It was found that the only effective way of countering this was to put the correct upright consistently on *S*'s nonpreferred side until the position habit was abandoned. In general, during Phase 1 a random position sequence was employed from which *E* departed periodically in order to avoid reinforcing incorrect hypotheses that *S* appeared to be developing.

At the beginning of the training in Phase 1 the uprights were placed on the table a few inches apart in front of *S*. They were gradually moved away from each other until one was on *S*'s extreme left and the other on his extreme right:

[10] The success criteria were selected following Hilgard (1951, Table 5, p. 535) which gives the number of successive correct responses required for statistical significance as a function of the total number of trials.

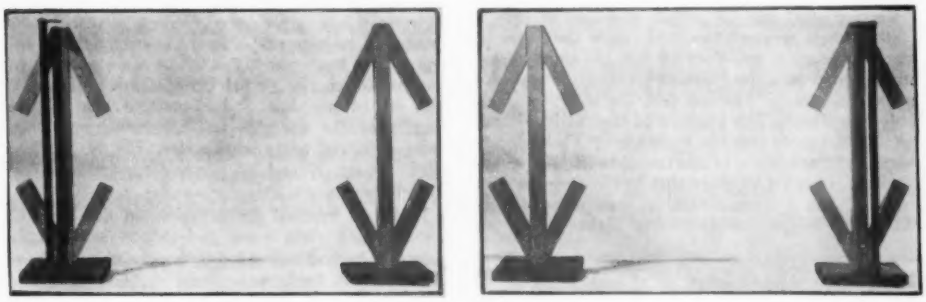

FIG. 3. Example of a measurement trial: the measuring stick is juxtaposed with each upright in turn.

these were the positions the uprights occupied in Phase 2, in order to render direct visual comparison of their heights more difficult.

Phase 2: Measurement and retention. This phase consisted of a minimum of 40 trials. Eighteen of them were "measurement" trials. Ten were "trials without yardstick" and there were at least 12 retention trials.

The uprights were always presented in pairs as in Phase 1, and the candy was always under the same upright (the "longer" or the "shorter") as it had been in Phase 1. In the retention trials the same combinations of stimuli were used as had been used in the training trials of Phase 1. However, in the measurement trials and trials without yardstick only uprights designated in Fig. 2 by capital letters were used. The objective differences in height between the stimuli were therefore always quite small. It will be noted that among the 11 uprights designated by capital letters in Fig. 2, there are 39[11] possible pairings in which there is a small true difference in height. In a few pretesting sessions the 28 pairs which seemed most difficult to discriminate by direct visual comparison were selected and 18 of these were randomly assigned to the measurement trials. The remaining 10 pairs were used in the trials without yardstick.

During measurement trials *E* juxtaposed one of the measuring sticks with each upright in turn, the left-hand one first, as shown in Fig. 3. The measuring stick was always shorter than one of the uprights and longer than the other. It was placed vertically on the base of each upright and rotated so that the spigot passed over the top of the shorter upright or rested against the side of the longer upright (see Fig. 3). If *S* did not seem to be watching during this procedure, the stick was knocked against the top of the upright to attract his attention and he was told, "Look at the tops of them." The *S* was not permitted to touch either of the stimuli until the measurement had been performed. The measuring operation was often prefaced by the words, "Wait, we are going to try and help you find the candy." If the child failed to lift one of the uprights immediately after the measuring operation had been performed, this operation was repeated and then he was instructed, "Find the candy." The criterion for successful "measurement" was 13 correct trials out of 18.

The trials without yardstick were the same as the measurement trials except that the entire measuring procedure was omitted. The *S* was instructed, "See if you can find the candy without help this time." The purpose of the trials without yardstick was to test the accuracy of *S*'s discrimination of the lengths of the uprights without measurement, in order to show that in the measurement trials *S* had "inferred" rather than directly perceived the longer (or shorter) upright.

[11] Fifteen pairings of A, B, C with P, Q, R, S, T + 9 pairings of A, B, C with X, Y, Z + 15 pairings of P, Q, R, S, T with X, Y, Z.

In the retention trials pairs of uprights clearly different in height, as in Phase 1, were used. The retention trials were included in order to ensure that *S*s who were unable to find the correct upright in the measurement trials still retained the concept "longer" (or "shorter"), i.e., to ensure that failure on measurement trials did not occur simply because *S* had forgotten the problem "instructions," taught in Phase 1. When failure occurred on any of the retention trials the original length discrimination was retaught as in Phase 1 to the same criterion of eight successive correct comparisons of eight different pairs of stimuli, before the measurement trials and trials without yardstick were continued. No *S* was retrained more than three times.

The measurement trials were given in pairs, and between every pair of measurement trials a trial without yardstick was inserted. The retention trials were distributed among the 28 measurement trials and trials without yardstick: between every fourth and fifth of these trials, two retention trials were inserted.

Throughout Phase 2 the position of the correct upright was varied randomly.

Phase 3. The procedure in this phase depended on whether *S* had reached the criterion of success in the measurement trials of Phase 2. If *S* had failed to reach this criterion, the following procedure was adopted. Two uprights similar in appearance (e.g., S and Z or T and Y in Fig. 2) were first placed next to each other and inclined together until the tops touched. The *S* was told to put his finger on the top of the one that was "taller or bigger" (or, "shorter or smaller"). The *S* then watched a piece of candy being put under the taller upright (or the shorter, whichever concept had originally been learned). The positions of the uprights on the table were then rapidly interchanged several times so that *S* lost track of the one with the candy. The uprights were then measured as in the Phase 2 measurement trials. As each upright was being measured, *S* was instructed to point to the "one that is longer or taller" (or, the "one that is smaller or shorter"). Thus, an *S*, for whom the longer upright had previously been rewarded, now pointed to the measuring rod as the shorter upright was being measured, and to the longer upright as this was measured. When both had been measured, he was told, "Find the candy under the one that is longer or taller." The main purpose of the supplementary verbal instructions in Phase 3 was to ensure that *S* had noticed the relative lengths of the uprights and measuring rod. Not more than six trials were given. To be credited with success, *S* had to respond correctly in five successive trials.

If *S* had reached the criterion on the Phase 2 measurement trials, it was desirable to make doubly sure that he had succeeded by a genuine process of inference. A book was placed under one of the pedestals so that the bases of the pedestals were no longer at the same level. Two uprights similar

in appearance (e.g., S and Z or T and Y) were again selected and placed one on each pedestal. The child was asked which was the taller (or shorter, if he had originally learned this discrimination). Following his answer the uprights were measured in the usual manner, and he was then asked whether his original judgment had been correct or not. The uprights were reshuffled several times and this procedure repeated until *S* had demonstrated without error in 5 trials that he was able both to maintain and to change his original judgment, depending on the facts as indicated by measurement.

Results and Discussion

Level of Performance Attained

One *S* failed to reach the learning criterion in Phase 1 and was therefore excluded from the subsequent phases.[12] The remaining 40 *S*s all learned the original length discrimination in Phase 1, and to all of these *S*s Phases 2 and 3 were therefore administered.

Fifteen *S*s failed to reach the criterion of 13 correct measurement trials in Phase 2, and also failed in Phase 3. These *S*s therefore failed to find the longer or shorter upright by inference in both phases. An additional 5 *S*s failed to reach the criterion in the measurement trials of Phase 2, but made an inferential response in Phase 3, when the nonverbal method was supplemented by verbal instructions. The remaining 20 *S*s were successful under both conditions; there were no *S*s who reached the criterion in Phase 2 but failed to do so in Phase 3.

The percentage of *S*s within each 8-mo. age interval between 3–6 and 6–10 who successfully found the correct upright following the measuring procedure is shown in Fig. 4. This graph shows separately the increase with age of the percentage of *S*s reaching the criterion in the measurement trials of Phase 2, and those who made an inferential response in either Phase 2 or Phase 3. It can be seen that a regular development appears to occur, and that in this sample the threshold age at which 50% of children made the inference studied is somewhere between 4–2 and 5–5. In Table 1 the development with age is shown in terms of the mean number of correct responses in the 18 measurement trials.

The improvement in performance with age in the trials without yardstick is also shown in Fig. 4 and Table 1. A criterion of eight correct choices in the 10 trials without

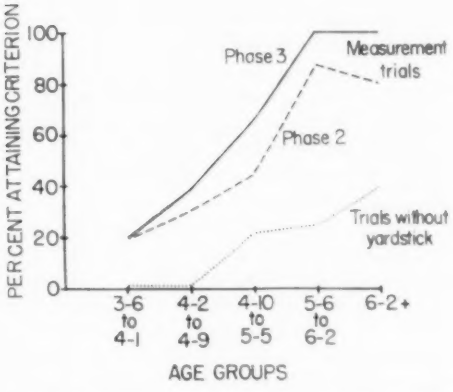

FIG. 4. Percentage of *S*s finding the correct upright by inference (measurement trials) and by direct comparison (trials without yardstick).

TABLE 1

NUMBER OF CORRECT RESPONSES AS A FUNCTION OF AGE IN THE MEASUREMENT TRIALS AND TRIALS WITHOUT YARDSTICK OF PHASE 2

Age Group	*N*	Mean No. Correct Responses	
		18 Measurement Trials	10 Trials Without Yardstick
3–6 to 4–1	5	9.5	5.0
4–2 to 4–9	13	10.2*	4.6
4–10 to 5–5	9[a]	12.4****	6.0
5–6 to 6–1	8	14.8****	5.7
6–2 to 7–0	5	15.6****	7.2**

Note.—The statistical tests were based on the null hypothesis that *S*s' scores are distributed as predicted by the binomial expansion, with N = no. of trials and $p = q = .5$. Thus, where n = no. of *S*s in an age group, z = (Mean Score of age group $- Np$) / $\sqrt{Npqn}$.

[a] One *S* excluded who failed to learn the initial length discrimination in Phase 1.

* Significant at the .05 level.

** Significant at the .01 level.

**** Significant at the .0001 level.

[12] This child was aged 5–0. In three training sessions she failed to learn even the first discrimination (between P and d). The writer has no explanation for this failure.

yardstick was used in Fig. 4.[13] It can be seen that in all except the oldest age group the average number of correct responses in the trials without yardstick is close to the chance level. Even in the oldest age group accuracy is considerably less than in the measurement trials.

Since the use of the measuring stick is the only factor that distinguishes the measurement trials from the trials without yardstick, the greater accuracy in the measurement trials is attributable to the *S*s' use of the cues provided by the measuring procedure. Further evidence that the children successful in the measurement trials found the longer or shorter upright by inference is provided by the fact that all these children were able to find the correct upright in Phase 3, when the bases of the uprights were at different levels.

The evidence that the children successful in the measurement trials were using an inferential procedure is thus convincing. The results suggest also that the lack of success of the children who did not reach the criterion in either the measurement trials of Phase 2 or in Phase 3 cannot be explained simply as a failure to retain the original concept "longer" or "shorter," taught in Phase 1. Evidence for this is found in the retention trials of Phase 2. Of the 15 children who failed in both the measurement trials of Phase 2 and in Phase 3, six responded correctly in all 12 retention trials, five failed in one retention trial and had to be retrained once, and three had to be retrained twice. One *S* was discontinued on Phase 2 because of consistent failure to retain the discrimination. Of the 20 children who reached the criterion on the measurement trials, three had to be retrained once and one twice. Thus, *S*s who failed to make an inferential response also tended to forget the discrimination taught in Phase 1 somewhat more often than *S*s who made an inferential response. This is not surprising since the former *S*s were in general younger and more distractible. But it is extremely unlikely that failure to reach the criterion in the measurement trials is due to a simple failure to retain the original discrimination. It may be recalled that the retention trials were spaced among the measurement trials and trials without yardstick in such a way that any failure to retain the discrimination would be very quickly discovered. If this happened, *S*s were then retrained before any further measurement trials or trials without yardstick were administered. Thus one or two failures of retention would not vitiate the results in the measurement trials of Phase 2. Moreover, an *S* is not considered to have failed to have made an inferential response unless he also failed in Phase 3 with the supplementary verbal instructions.

Three variables probably unrelated to successful "measurement." Whether or not an inferential procedure was used seemed unrelated to the location of the candy under the longer or shorter upright. For 20 *S*s the reward was consistently under the shorter upright, and of these, eight failed to make an inferential response. Of the remaining 20 *S*s, seven failed to find the longer upright by inference.

Sex of *S* also seemed to be unrelated to success or failure in finding the correct upright by inference. Of 18 boys, eight failed to make an inferential response, and of 22 girls, seven failed. The difference is statistically insignificant.

Some of the *S*s received the order-discrimination task described in the next section before receiving this task. Of the total group of 40 *S*s, 16 received the order task first. Of these only four failed to make an inferential response, whereas 11 of the remaining 24 *S*s failed. However, the *S*s who received the order-discrimination task first were generally older. This was because a few young *S*s inopportunely went on vacation after receiving the order task. Unlimited time was not available and it was therefore desirable to begin the measurement task early, since, owing to the lengthy Phase 1 training, it required several ses-

[13] This criterion errs on the lenient side: $p = .054$ of obtaining by chance eight or more correct responses in 10 trials. This criterion was selected because it is comparable to the criterion used in the measurement trials ($p = .048$ of finding 13 or more correct in 18 trials).

sions to administer to the younger children. Of the 21 children under 5-0, 14 took the measurement task first and 10 failed to make an inferential response. Of the seven who took the order task first, four failed to make an inferential response. Chi square, even without Yates correction, is very small; even for the whole group of 40 *S*s chi square is not significant. Thus there is no evidence that prior experience with the order task affected the *S*s' success in making an inferential response in the measurement task.

Temporal course of performance in Phase 2. In Fig. 5 the temporal course of performance in the 18 measurement trials is shown separately for the group of 20 *S*s who reached criterion and for the remainder who did not. It can be observed that there was no tendency to learn on the part of *S*s who failed; their performance never deviated far from the chance level. The 20 *S*s who reached criterion performed significantly better than chance even on the first three trials, and the improvement that followed took place very quickly. Moreover, even on the first measurement trial, 16 of these *S*s found the correct upright.[14] The evidence therefore indicates that *S*s who reached criterion on the measurement trials tended to use the inferential procedure from the beginning of Phase 2.

In Fig. 6 the course of performance on the 10 trials without yardstick is shown. It is noteworthy that the *S*s who reached the criterion on the measurement trials performed consistently better than the remainder. That is, the *S*s who made an inferential response also tended to discriminate the relative lengths of the uprights better when no inferential procedure was involved. In the last four trials without yardstick their performance even rose to a level significantly better than chance.

This correlation raises some very interesting questions. In particular, it might lead one to suggest that the dichotomy originally made between "inferred" and "perceived" length differences may not be as absolute psychologically as it is logically. This question is complicated and a discussion of it is deferred until it can be treated within the context of the totality of the results.[15] At this point one may note only that the *S*s who made an inferential response were older. Also, in the specific stimulus conditions used in this experiment, discrimination of the lengths of the uprights without the aid of the measuring stick is quite complex, because of the slight degree of illusion

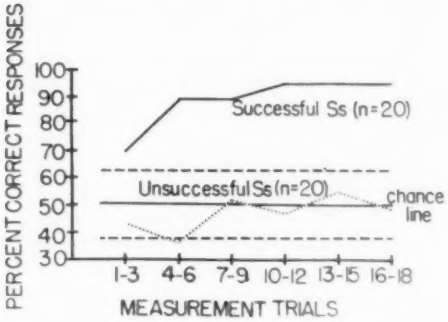

FIG. 5. Temporal course of performance on the measurement trials. The "successful" *S*s are those who reached the criterion. The parallel broken lines enclose the region between the minimum score significantly greater than chance and the maximum score significantly less than chance ($p = .05$).

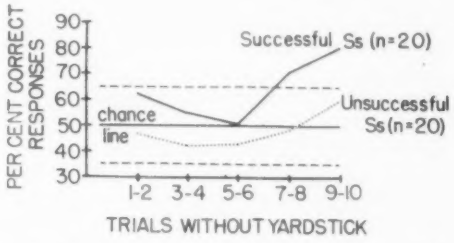

FIG. 6. Temporal course of performance on the trials without yardstick. The "successful" *S*s are those who reached the criterion in the measurement trials. The parallel broken lines enclose the region between the minimum score significantly greater than chance and the maximum score significantly less than chance ($p = .05$).

[14] $p < .05$ that this is a chance occurrence.

[15] See section entitled Changes in "Set" Associated With the Emergence of Measurement.

presumably induced by the arms at the upper and lower extremities of the uprights.[16]

Qualitative results. For the purposes of description, the *S*s' behavior may be grossly categorized into three types. However, there were many *S*s who did not fall clearly into any type. Also, at the beginning of the experiment it was not known how they would behave. On looking back at the protocols, it was found that for some of the *S*s run first there was no adequate record of the presence or absence of many of the behaviors later observed carefully.

The *S*s who failed to make an inferential response appeared to have no comprehension of the measuring operation in Phase 2. They did not spontaneously watch the measuring procedure except on the first few trials. Their attention had to be constantly attracted to the measuring sticks by various means, with the result that administering Phase 2 so that the measurement was at least observed, if not understood, proved to be quite exhausting to *E* and probably to *S* also, and often required several sessions. On the trials without yardstick these *S*s did not seem to compare the uprights visually. They gave no more than a quick glance at each and then reached for one of them. It was as if the possibility of an imperceptible difference in length was beyond their comprehension.

To others, who were generally the older *S*s, the nature of the task was quite clear. For example, one *S* verbalized, "I know one's bigger but they look the same." These *S*s would watch the measurement intently. During the trials without yardstick they attempted to assess the uprights carefully and often leaned back as far as they could in their chair to bring them closer together in their field of vision; they often glanced from one upright to the other several times. In general, these *S*s found the measurement trials rather easy and one even urged *E* not to use the measuring sticks because that "gave it away."

There was a very interesting third group, into which fell many of the younger *S*s who made an inferential response. Like the first group, they did not spontaneously watch the measuring procedure in Phase 2, but when their attention was drawn to it, they proved able to find the candy with the aid of the measuring stick. Some of these *S*s might find the candy two or three trials in a row, while paying attention, and still fail to watch spontaneously the measuring procedure in subsequent trials. Their behavior recalls the observations of Goldstein and Scheerer (1941) in which they argue that abstract thinking is an "attitude" rather than an "ability"; these *S*s were able to make an inferential response, but did not spontaneously attempt to do so. Some of the *S*s in this group reached the criterion in the measurement trials of Phase 2, whereas others only demonstrated their inferential ability in Phase 3, where the verbal instructions permitted *E* to take a more active role in fixing the *S*s' attention. Like the *S*s of the first group, these *S*s showed no comparison behavior in the trials without yardstick of Phase 2. This group was small: only eight *S*s fell unequivocally within it.

Role of verbalization. The spontaneous verbalizations of the children did not throw much light on the processes involved in solving the task. Of the task-related verbalizations, most occurred during Phase 1 or during the retention trials of Phase 2 (i.e., when the difference in length between the uprights was large). Examples of these are: "It's under the big one," "The little one," "The baby one," "Which one's up," "The teeny one." Interestingly, it appeared that when an *S* says, "It's under the big one," this should not necessarily be construed by *E* to mean "It's always under the big one." An *S*'s verbalization often seemed to refer only to a particular trial, and he might select the incorrect stimulus on the very next trial.

There was no evidence that verbal cues were of much help in solving the measure-

[16] It will be recalled that the uprights were deliberately constructed in this manner in order to increase the difficulty of discriminating their relative lengths without the aid of the measuring sticks. If the *S*s had shown a moderate or high degree of success in the trials without yardstick, it would have been difficult to prove statistically that *S*s who were successful in the measurement trials had used an inferential procedure.

ment trials. Spontaneous verbalization during Phase 1 of the location of the candy under the longer or shorter upright was not related to the performance of an inferential response. No child spontaneously verbalized the measuring procedure. At the end of the task several *S*s were asked how the use of the measuring stick helped them find the candy. All except two said they did not know. Of these two, one answer was, "Because it [the stick] makes it [the upright] smaller." The other answer was, "If the thing [measuring stick] is bigger than there [upright] and the other one the stick's smaller, it's there." Both these *S*s were girls aged 5-6. The latter child was the one, cited earlier, who verbalized, "I know one's bigger but they look the same." The performance of this *S* was interesting. During the measurement trials she pointed occasionally to one or the other of the two uprights, as if she were trying to verbalize the solution of the problem to herself subvocally. These efforts seemed to be more confusing than helpful to her: in the 18 measurement trials, she failed the first trial, solved the next six, then made four successive errors, followed by seven successive correct responses. She did not become similarly confused in the trials without yardstick, finding the objectively taller upright after careful visual comparison in 9 of the 10 trials. Even when the question "Which one is taller or bigger?" was posed verbally by *E* in Phase 3, she almost became confused again but managed to correct herself. She gave the impression that she was trying to form the series long-upright—measuring stick—short-upright, but, having formed the series, she forgot which end was which!

In general, although the location of the candy under the longer or shorter object was often verbalized, there was no evidence that any verbal cues were used which related specifically to the inferential procedure: no audible words were used in the measurement trials which were not also used in the initial learning and in the trials without yardstick. In the one *S* who appeared to be seeking to use verbal cues during the measuring procedure, these cues seem to have been more confusing than helpful.

Significance of Results for Piaget's Theory

The most important single result of this study is the finding that the inferential response $A > B.B > C. \supset .A > C$ can be elicited in children substantially earlier than Piaget leads one to believe. If Piaget's finding that children do not manipulate measuring instruments effectively until around age 7 or 8 yr. is valid, then children can perform this inferential response long before they are able to measure adequately. While the availability of this response may be a necessary condition for adequate measurement, it is therefore not a sufficient condition. One plausible reason for the apparent time lag between the development of the inference studied in this experiment and the development of adequate measuring behavior in Piaget's research is that manipulation of measuring instruments requires considerable learning of specific measuring techniques. Adequate measuring behavior may also demand interest in measurement and in achieving the additional accuracy that it makes possible. It is likely, therefore, that many of Piaget's studies, which he believes investigate the development of children's performance of the logical operations implicit in measurement, should be reinterpreted as studies of the development of skill and interest in the techniques of measurement. Inferential responses of the kind investigated here may be available to many of his *S*s, particularly those of intermediate age.

The experimental technique used also differed from Piaget's in the manner in which the stimuli were presented. Piaget used instructions which contained the words "same length," and "measure," or synonyms of these. Until Phase 3 at least, the task was presented nonverbally, except for the instruction, "Find the candy"; moreover, the words "longer" or "shorter," introduced in Phase 3, were preceded by the considerable training in length-discrimination involved in the previous phases.

Probably this difference in technique is a factor contributing to the difference between the present results and Piaget's. It is particularly likely that the nonverbal method facilitated the success of the youngest children making an inferential response.

The fact that the findings of this experiment indicate that Piaget is mistaken about the specific age at which the inferences basic to measurement develop raises the question: Is it possible—and if possible, is it desirable—to attempt to decide at what age these inferences develop in children? In this sample of children, presumably of average intelligence, and under the particular stimulus and response conditions employed here, this age was found to be somewhere between 4-2 and 5-5. However, it would be incautious to assert that this is the earliest age at which these inferences can be elicited in average children. Although a considerable effort was made to use experimental operations simpler than Piaget's, it is quite possible that further study would lead to the discovery of procedures in which the inferential response studied could be elicited in still younger children, thus demonstrating that some relevant factor was overlooked in the present experimental design.

The crucial theoretical issue concerns the usefulness of postulating the existence of an "emergent" formal reasoning factor in the development of children's length judgment.

It could be argued that if the apparent "age of emergence" is a function of the particular manipulanda and discriminanda involved and of their mode of presentation to *S*, then the concept of an emergent level is unsatisfactory in that it lacks parsimony and could even obstruct psychological analysis. One might be forced to postulate as many "levels" as there are different experimental operations for eliciting the same alleged psychological process.

One weakness of this argument against the "levels" concept is that it could clearly be transposed and used to attack, not only Piaget's, but any concept of emergent levels. Moreover, one can argue that there must be some point between birth and maturity at which the kind of inferential process studied in this section develops in the average child. The reason why Piaget's results differ from those reported here is not simply that there is a difference in experimental procedures, but is, the writer believes, because in designing his experiments Piaget fails to eliminate important variables which are not involved in the definition of the processes he sets out to investigate. That is to say, in his studies of measurement, Piaget claims to be studying the development of the formal reasoning inherent in measurement; he is probably actually studying primarily the development of skill or interest in the use of measuring rods, or perhaps the development of children's understanding of certain words. The results obtained here therefore lend support to the notion of an "emergent" reasoning process in the development of children's length judgments, although the establishment of "emergent levels" is a more complex experimental problem than Piaget apparently seems to believe.

It is clear that if one seeks to state an age at which a particular type of response develops, the only age which is not completely arbitrary is the earliest age at which this type of response can be elicited using the simplest experimental procedures. To establish these "simplest" procedures requires the discovery and systematic exploration of all relevant factors. The experiment reported can be interpreted as a step in this direction. Such a program of research would presumably result either in a clear operational definition of the alleged "emergent" process, or in convincing evidence that the notion of an "emergent" reasoning process in children's conception of length is unsatisfactory because no clear distinction between the "higher" and "lower" processes can be found. There is nothing in this study to suggest that the formal reasoning inherent in measurement cannot be defined experimentally.

Changes in "Set" Associated with the Emergence of Measurement

In presenting the qualitative results, differences in behavior were described, which

seemed broadly related to age. The oldest *S*s showed the most comparison behavior, especially in the trials without yardstick: they tended to lean back in their chairs to facilitate visual comparison of the uprights; they paused before responding and looked intently at each upright, often giving each several glances. Such behavior indicates that they readily transferred the set to discriminate length, learned in Phase 1 with the large and obvious length differences, to Phase 2 where the length differences were small. In contrast to this, comparison behavior was conspicuous by its absence in many of the youngest *S*s. This fact suggests that perhaps differences in length too small to be readily perceptible did not have much significance for the younger children. It was as if they were not "trying" to find the longer upright when the differences were small, not because they were unmotivated, but because they failed to recognize the possibility of a difference in length too small to be noticed. They seemed to lack the attitude of the older *S*s expressed in the words, "I know one's bigger, but they look the same." [17]

Thus, there is evidence that two kinds of behavior develop. First, there is the development of the inferential response, which is revealed quantitatively in the increasing success with age in the measurement trials of Phase 2. Second, there is a change with age in the "attitude" or "set" adopted toward small differences in length, which seems to consist in a greater tendency with increasing age to transfer the set to discriminate length from large obvious differences to small barely perceptible ones. This change in "attitude" is revealed quantitatively in the better-than-chance success of the oldest *S*s in the trials without yardstick.

The question must be raised whether there is any connection between these two phenomena. Is the development of the inferential response related to the phenomenon referred to as a change in attitude? On this issue there are three possible hypotheses which are worth exploring, although the present data do not clearly indicate which hypothesis is valid.

The first hypothesis is that the change in "attitude" or "set" and the inferential response develop quite independently and are unrelated in the individual child. If this is so then there should have been no tendency for the *S*s who made an inferential response to have done better in the trials without yardstick than those who failed to make an inferential response. In the total sample this was certainly not true. The relationship can be clearly seen in Fig. 6. However, when age was taken into account the correlation was equivocal. Matching the seven youngest *S*s who made an inferential response with *S*s of the same age who failed to do so, it was found that those who made an inferential response obtained a mean score on the trials without yardstick significantly higher than the others ($t = 2.47$, $df = 6$, $p < .05$); however, both means (5.7 and 4.4) can themselves be readily explained on the hypothesis that the *S*s were responding at random. Consequently the evidence against the first hypothesis is ambiguous. The remaining hypotheses both assume a relationship.

The second hypothesis is that the change in attitude is the determining variable and that the development of an inferential response is dependent on it. That is, the adoption by *S* of the more "mature" set is a precondition for the occurrence of an inferential response: unless *S* realizes the uprights differ in length and consequently seeks the longer, he will not infer the longer upright in the measurement trials. Plausible though this hypothesis may be, the evidence is against it. The best evidence is qualitative. Although all *S*s who showed comparison behavior also made an inferential response, there were at least eight *S*s who made an inferential response but who showed no comparison behavior in the trials without yardstick: there was no

[17] If this interpretation is correct, it raises the more general problem of how and when children learn to make the conceptual distinction between a "real" and a "phenomenal" world. This problem has been largely ignored by experimental genetic psychology outside Geneva. No general discussion of it is attempted here.

change in posture to bring the uprights closer together in the field of view, the response was not preceded by several ocular shifts from one stimulus to the other, and there was no verbalization which suggested that these *S*s realized the stimuli differed in length. Thus, an inferential response could occur in the context of behavior associated with the more "primitive" set. The quantitative evidence of Fig. 4 points in the same direction: inferential responses occurred in much younger *S*s than did any success in the trials without yardstick. Unfortunately, however, the quantitative evidence is inconclusive because of the illusion presumably induced by the arms placed at the upper and lower extremities of the uprights: the possibility cannot be dismissed that the younger *S*s failed to discriminate to a degree better than chance because they were deceived by the illusion, and not because their "set" was inappropriate.

The evidence, although primarily qualitative, is more consistent with the third hypothesis, which is that the change in attitude toward barely perceptible length differences is dependent upon the prior development of inferential processes of the type studied. Should this hypothesis prove valid, it could be explained on the theory that the development of the inferences basic to measurement leads to the establishment of a concept of length as a dimension in which continuous small increments are possible. A frame of reference might then be created within which barely perceptible length differences could acquire meaning.

To decide definitively between these hypotheses a detailed exploration would be necessary of the minimum length difference discriminated as a function of age, both with and without measurement, and with stimuli which do not give rise to an illusion. It would be interesting to know whether the smallest length difference noticed by children decreases steadily as a function of age, or whether there is a sharp decrease at any period which might correspond to the change in "attitude" toward small differences which is suggested by the present data.[18] If there is such a sharp decrease, it would be interesting to know whether or not it is associated with the development of inferential responses of the kind studied here, and whether these inferential responses usher in the sharp decrease in DL, or follow it. Some of the issues raised by this study might then be resolved.

Recapitulation

Some aspects of the logic of length measurement were discussed and a series of studies of length judgment were reviewed. These included a group of studies by Piaget in which he attempted to diagnose children's grasp of the logic of measurement through their manner of using measuring rods. He found children under 7 yr. deficient in this respect. A number of criticisms of Piaget's work and general methodology were discussed. It was suggested that adequate use of measuring rods by children may involve factors other than the performance of the logical operations basic to measurement. In particular, learning of effective measuring techniques may be an important factor.

To investigate this possibility a task was constructed in which no manipulation of measuring rods was demanded of the *S*s. They were required to make deductions about the relative lengths of two nearly equal upright sticks, using the rule, $A > B.B > C. \supset .A > C$. This inference has the status of an axiom in the logic of measurement. The experimental procedures were adapted from work with animals, so that it was possible to present the task with a minimum of verbal communication between *E* and *S*. Vocabulary development, an uncontrolled factor in Piaget's studies, is therefore unlikely to be a factor influencing the results. The methodological importance of this innovation was stressed.

[18] It is assumed, of course, that the DL is explored using an experimental technique in which *S* is first trained to discriminate large length differences, and then the smallest length difference to which he will generalize this discrimination is studied.

It was found that with these procedures, the inferential response A > B.B > C. ⊃ .A > C was elicited approximately two or three years before the age at which Piaget claims it first becomes available to children.

Unless Piaget incorrectly reports the age at which children begin to manipulate measuring rods properly, it is clear that factors other than correct reasoning according to the axiomatic standards of the logic of measurement are involved in his studies. These factors are probably vocabulary development (particularly, comprehension of the meanings of words such as "measure," "same length," etc.) and the development of skill and interest in the technique of length measurement. The effect of these factors is probably to conceal the reasoning ability of many of Piaget's *S*s.

The results indicate that the definition of "emergent levels" of development is a complicated problem. The objection that the apparent "age of emergence" is dependent on the specific experimental procedures utilized does not pose insuperable difficulties for the concept of "levels." This objection does, however, suffice to demonstrate that any stated "age of emergence" must be arbitrary until the effects of all relevant variables have been systematically explored, so that extraneous factors (e.g., vocabulary development, skill in the use of measuring rods, etc.) which may hinder the younger *S*s from responding to the relevant cues (the relative lengths of uprights and measuring stick) are identified and eliminated. With this qualification, the experiment reported lends support to Piaget's notion that the formal reasoning inherent in length measurement constitutes an "emergent" process in child development.

The experimental task also contained some trials in which the two nearly equal uprights were compared by direct visual comparison. Qualitative aspects of the behavior of some of the older *S*s in these trials suggested that they adopted an "attitude" or "set" which was different from that of many of the younger *S*s, and which assisted them to achieve some measure of success in these discrimination trials. The possible genetic relationship of this change with age in "attitude" or "set" to the development of the inferential response was discussed. The evidence, although not conclusive, suggests that once a child is capable of the inferences basic to measurement, he may develop a concept of length as a dimension, as a result of which he more readily adopts a "set" to discriminate small differences in length. A method of examining this hypothesis experimentally was suggested.

Development of a Concept of Position Order

In this section Piaget's views on the development of concepts of order are examined in a way which parallels the previous examination of his views on the development of length measurement. A study of his is described and an experiment is reported which investigates certain task parameters unexplored by Piaget.

Piaget's Study

In the Piaget and Inhelder (1956, Ch. 3) study, seven to nine beads of different colors threaded on a wire served as a model which *S* had to copy by choosing similar colored beads and threading them in the same order on another wire. For some of the *S*s, instead of beads on a wire, dolls' garments differing in color or form or both were hung on a toy clothes line, and *S* was told to choose similar garments and to hang them in the same order on another clothes line. The *S*s' ability to distinguish among the beads and garments was established before the experiment.

Piaget required his *S*s to copy the order of a linear model and also to make various transformations, as demanded by the following spatial arrangements of model and copy: 1. model linear, copy linear and directly beneath the model; 2. model linear, copy linear but displaced to one side; 3. model circular, copy linear; 4. model linear, copy linear with order inversed; 5. model in the form of a "figure 8," copy linear.

Piaget does not discuss how he conveyed the instructions to the *S*s. This must have been a problem since *S* need not interpret the word "same" to mean "same order," and any instruction which contains the word "order" presumably will only be understood by a child who already possesses a concept of order to which he can assimilate this word. Neither does Piaget mention how he specified the starting position to the child when the model was nonlinear.

Piaget does not publish any tabulated results. He says the responses could be divided into three stages (with substages) of development between 3 and 7 yr. In Stage I, from about 3 to 4½ yr. of age, *S* was able to select similar beads or garments but was not able to reproduce accurately the order prevailing in the model. In Stage IIA, the model was correctly reproduced only when linear and directly above the copy, so that careful matching of each bead with the one immediately above it was possible. In Stage IIB, the copy was correct when displaced to one side of a linear model, and also when the model was circular, but failure still occurred on the other two arrangements, indicating to Piaget that the notion of position order was not yet completely abstracted. In Stage III, beginning at about age 7 yr., all arrangements were accurately copied.

Whether the material to be ordered consisted of colored beads or of dolls' garments was apparently immaterial.

It may be noted that Piaget takes it for granted that the five settings he used are on a continuum of difficulty. Placing them all within the same logical category of transformations of ordered sequences, he assumes that the five settings represent stages of development of the same psychological process. However, it is possible that there might be psychological processes or abilities involved in transforming a cyclic order or "figure 8" into a linear order, which are not involved in a linear-to-linear translation. This could only be investigated through an examination of the intercorrelations of success and failure between the settings (not reported by Piaget), or by a longitudinal study of the development of individual children through the stages.

In some of the omnibus intelligence tests there are individual items in which success appears to demand some concept of order. For example, on Form M of the Binet there is an item which demands copying a chain of beads, which remains exposed. Test standardization (Terman & Merrill, 1937) places this test at Year VI. No disagreement with Piaget's findings is suggested.

In tasks which require the copying or matching of sequences, such as those described, there are three obvious parameters. The formal arrangement of the stimulus (linear order, inverse linear, cyclic, etc.) is one such parameter. This is the one varied by Piaget.

The number of objects to be ordered is another obvious parameter. One might reasonably expect that a sequence of three objects would prove less difficult than sequences of five or seven or more. This parameter is studied in detail in the experiment reported below.

The third obvious parameter is the nature of the material ordered. Although Piaget found that copying a sequence of dolls' garments did not differ in difficulty from copying a sequence of differently colored beads, this parameter has not been studied systematically. Two possible content parameters are the degree of familiarity of the material and its attractiveness to the child. Examination of content parameters is important, since if order concepts are entirely content-specific, i.e., if as many concepts of order develop in children as there are objects capable of being arranged in sequence, then one would have to question the usefulness of Piaget's postulation of the development of a general concept of position order.

In addition to the above parameters, there may be other factors involved. It is possible to devise a number of different experimental procedures for eliciting order concepts, and it is likely that the age at which order concepts will be found to develop will vary with the specific experimental procedures selected. In Piaget's study the instructions were communicated verbally

and the child's response was to copy a model provided by *E*. The adequacy of the verbal communication between *E* and *S* therefore becomes a factor of unknown importance.

In a single experiment it is impossible to study the effects of varying all the above parameters. In the study reported below the parameter "number of objects ordered" has been singled out for special investigation. While recognizing that content parameters may later prove important, it seemed that the parameter selected should be studied first, and the material remained the same throughout the experiment. It consisted of colored discs pasted on cards and only the number of colors was varied. The experimental procedures were developed from nonverbal techniques used in work with animals. The "matching-from-sample" technique was employed. There were few verbal instructions, and the words "same" and "order" were eliminated. This use of relatively nonverbal methods constitutes an important departure from Piaget's experimental technique.[19] All stimulus settings were linear; consequently the ages at which order concepts were elicited are not necessarily generalizable to stimulus settings which are nonlinear, like some of those used by Piaget.

Method

Subjects

The *S*s were 46 children from the nurseries described in the report of the experiment in the previous section. They ranged in age from 3-6 to 7-3. The distribution by age is shown in Table 3. There were 20 boys and 26 girls. Thirty-seven *S*s also received the measurement task described earlier, and of these, 21 took the measurement task before being given the order-discrimination task.

Materials and Procedure

The task consisted of a series of problems in which *S* was required to find an item identical to a sample item. The task began with simple color matching: *S* had to find that one of three discs which was the same color as a sample disc. In subsequent problems *S* had to match color sequences. In the initial series of these the sequence was only two units long, e.g., *S* had to match a sequence in the order red-gray (say) against one of two sequences, gray-red and red-gray. In the next series of problems the sequence had three units, e.g., *S* matched a sample in the order red-gray-green (say) against one of three sequences, gray-red-green, red-green-gray, and red-gray-green. In the fourth and fifth series of problems the sequences were five and seven units long, respectively.

Materials. These consisted of blank filing cards on which colored paper discs 1 in. in diameter had been pasted. The colors used were red, yellow, green, black, walnut, Prussian blue, and a light sky-blue gray. There were five sets of cards. Set A consisted of 14 cards with a single disc in the center of each. Set B consisted of 21 cards, with a pair of unlike colors on each card. Seven pairs of colors were selected and each pair was represented three times. Set C consisted of three sets of six cards with three colors on each card. In Set D there were two sets of six cards with five colors on each card. Set E comprised 10 cards with all seven colors on each card.

In each of Sets B, C, D, and E, the sequence of colors on any card could be reversed by inverting the card. The precise sequences of the colors on the cards in each trial is shown in Table 2. It can be seen that in Sets D and E the first and last colors were the same on all cards in each trial so that *S* could not match the sample without paying attention to the internal order of the colors.

Procedure. The *S* sat facing the short edge of a low rectangular table and *E* sat on his right facing the long edge. The sample card was placed flat on the table immediately in front of *S* and the choice cards (i.e., the cards from which *S* was to choose the one matching the sample) were placed equidistant from him about 2½ ft. away on the table, where they could easily be manipulated by *E*. Placing the choice cards beyond *S*'s convenient reach was found to be the best procedure. Being farther away, the choice cards were closer together in his field of view, thus encouraging comparison behavior instead of an unthinking dive for one of the cards. This arrangement also prevented *S* from attempting to reach for more than one card at once, thus removing a possible source of discord between *E* and *S*!

The choice cards rested against small white open cardboard boxes placed face down on the table. The cards were thus supported at an angle 30°–45° with the table surface. A small piece of candy was placed under the box which supported the card which duplicated the sample. Between trials a cardboard screen placed on the table hid the boxes from *S*, and *E* was careful that no noise or other cues occurred as the candy was placed in position. There were always three choice

[19] See the discussion of this point in the previous section.

TABLE 2

COLOR SEQUENCES USED IN SERIES B, C, D, AND E OF THE ORDER-DISCRIMINATION TASK

Trial	Sample	Choice Cards
	Series B	
1–6	RS	RS and SR
7–12	SR	RS and SR
13	PY	PY and YP
14	YP	PY and YP
15–26	as Trials 13–14, using BS and SB, GW and WG, etc.	
	Series C	
1	SGR	SGR, GSR, and SRG
2	GSR	SGR, GSR, and SRG
3	SRG	SGR, GSR, and SRG
4–6	as Trials 1–3, using R, B, and Y	
7–9	as Trials 1–3, using Y, W, and P	
10–18	as Trials 1–9, cards inverted	
	Series D	
1	RPYGW	RPYGW, RPGYW, and RYPGW
2	RPGYW	RPYGW, RPGYW, and RYPGW
3	RYPGW	RPYGW, RPGYW, and RYPGW
4–6	as Trials 1–3, using Y, R, B, S, and G	
7–12	as Trials 1–6, cards inverted	
	Series E	
1	SWBGYRP	SWBGYRP, SBWGYRP, and SWBGRYP
2	SBWGYRP	SWBGYRP, SBWGYRP, and SWBGRYP
3	SWBGRYP	SWBGYRP, SBWGYRP, and SWBGRYP
4	SWBYGRP	SWBYGRP, SWGBYRP, and SWBGYRP
5	SWGBYRP	SWBYGRP, SWGBYRP, and SWBGYRP
6–8	as Trials 1–3, cards inverted	

Note.—B = black, G = green, R = red, Y = yellow, W = walnut, S = sky-blue gray, P = Prussian blue.

cards, except when the cards of Set B were used, when there were only two. The position of the correct choice card was varied randomly, with the restriction that the same position was never correct in more than two successive trials. As the sample was presented to *S*, *E* asked, "Where is this one?" or said, "Now find this card." *S* indicated his response by pointing and was given the candy if correct. If the pointing was equivocal, the boxes were moved just within *S*'s reach so that he could find the candy himself.

The following "training" procedure was adopted in all trials, except in the first 12 trials of Series B. If *S*'s response was incorrect, he was first shown that there was no candy under the box to which he was pointing. Then the sample was held over the correct card with the words, "Look, here it is," and the box was lifted to expose the candy which, however, was not given to *S*. When *S*'s response was correct, the sample was held over the indicated card and *S* was told, "That's right." The box was lifted and the candy given to *S*.

Five series of trials were given, each series employing the cards of one of the sets. With Series A which was presented first, each *S* was taught to match from sample. This series was discontinued following six consecutive correct matchings. Series B began with six trials in which the sample was a sequence red-gray, and the choice cards were gray-red and red-gray. In the first four of these trials, after the sample had rested in front of *S* for a few seconds, it was picked up and held over each of the choice cards in turn for about 10 sec. to facilitate visual comparison of the sample and choice cards. The *S* was asked, "Is it this one? Or is it this one?" Any pointing by *S* during this period was ignored. The last two of the six trials followed the regular procedure. These six trials using the sample

sequence red-gray were followed by a similar six trials with the sample gray-red, and the same choice cards. As shown in Table 2, there then ensued 14 generalization trials in which each sequence in Set B served as sample once. The *S* was considered successful in Series B if he met either of the following criteria: (*a*) correct matching in all of the first five generalization trials, in which case Series B was discontinued, or (*b*) correct matching in 11 of the 14 generalization trials. All *S*s who failed to meet either of these criteria went through all 14 trials.

Series C was administered whether or not the criteria for success on Series B had been met. Series C consisted of 18 trials, all of which were treated as generalization trials. These trials are shown in Table 2. The *S* was considered successful in Series C either following five consecutive matchings, or if correct matching occurred in 11 of the 18 trials.

Series D was administered unless *S* had failed to meet any of the criteria for success in either Series B or C.[20] Series D consisted of 12 trials, as shown in Table 2. Criteria for success were either five consecutive correct responses, or a total of 8 out of 12 trials correct. If either of these criteria was met, Series E, with seven colors per card, was administered. This series consisted of eight trials with passing criterion six trials correct.

The criteria set for success in the five series were selected so that one would expect them to be attained by approximately 5% of *S*s who responded randomly.

Administration of the entire order-discrimination task often required more than one session, although rarely more than two sessions. At the beginning of all sessions subsequent to the first, Series A was re-administered.

Results and Discussion

Level of performance attained. With the exception of two boys who were therefore excluded from the remaining four series, all 46 *S*s readily learned to match from sample when there was only one color on each card (Series A). Twenty-two *S*s met one or another of the criteria for success on Series B and 22 failed; on Series C, 29 *S*s met a criterion, 15 failing. The 22 *S*s who succeeded on Series B also did so on Series C. With Series D there were 25 successes and with Series E 11 successes.

The development with age of success on each series is shown in Table 3. It can be seen that no child under 4-2 showed any concept of order at all. Surprisingly, significantly fewer children matched the order of sequences of two items than the order of sequences of three items.[21] Order was apparently most readily discriminated when there were three colors on the cards.

One possible explanation for the differing incidence of successful matching on Series B and C was examined. It was thought that in responding to Series C, the children might have profited from the training provided by the trials of Series B, even though they failed to match on these trials. This explanation implied that if some of the trials of Series B were repeated following the administration of the later series, matching of order would be found on Series B. To explore this possibility, four of the seven *S*s who failed Series B but solved Series C were given between 5 and 10 extra trials on Series B again after the longer series had been administered. All four of these *S*s made several errors, indicating that

TABLE 3

PERCENTAGE *S*s IN EACH AGE GROUP REACHING CRITERIA FOR SUCCESS IN THE SERIES OF THE ORDER-DISCRIMINATION TASK

Age Group	*N*	No. of Colors in Sequence			
		2	3	5	7
3–6 to 4–2	7[a]	0	0	0	0
4–3 to 4–11	12[a]	25	50	33	0
5–0 to 5–8	15	67	87	73	27
5–9 to 6–5	6	83	100	100	67
6–6 to 7–3	4	100	100	100	75

[a] One *S*, who failed to learn to match-from-sample with one color per card, is excluded.

[20] It seemed unlikely that a child who had failed to discriminate order with either two or three colors per card would suddenly begin to succeed when cards containing five or seven colors were introduced. Moreover, while the 12 trials of Series D were sufficient to demonstrate generalization of order discrimination from two or three colors per card to five colors, it would be quite unfair to expect *S*s who had not demonstrated order discrimination in the 44 trials of the previous two series suddenly to demonstrate this within the mere 12 trials of Series D.

[21] Chi square (correlated proportions, $df = 1$) = 7.1 ($p < .01$). Sequences of two and five items did not differ significantly in the number of *S*s solving them.

practice with the longer series had not caused any striking improvement in performance on Series B.

This seemed to leave no good explanation for the greater difficulty of Series B over Series C. The children's verbalizations provided no clues; there were in fact practically no task-related verbalizations during the entire task. No particular differences in eye movements were noticed during responses to Series B and C. Latency of response was not noticeably greater on Series B than on Series C, although in general children who matched successfully appeared to take somewhat longer to respond on all series. An interesting possibility [22] is that Series B may have been more difficult because the order of the colors on the wrong choice card is a symmetrical reversal of the sample sequence on this series, whereas in none of the subsequent series do the choice cards contain an inversion of the sample.

There is no evidence that success in discriminating order is affected either by *S*'s sex or by prior experience with the measurement task. Of 18 boys, 56%, 61%, 56%, and 22% succeeded on Series B, C, D, and E, respectively. The corresponding figures for the 26 girls were 46%, 69%, 58%, and 30%. The differences between boys and girls are all statistically insignificant.

Of the 44 *S*s, 21 had prior experience with the measurement task. Of these, nine failed to show any concept of order. Of the 23 *S*s without prior experience with the measurement task, six failed to show any concept of order. This difference is not statistically significant. The paradoxical direction of the difference may be accounted for by the fact that, as explained earlier, a majority of the younger *S*s were given the measurement task first.

Significance of results for Piaget's theory. In his experiment reviewed above, Piaget used sequences which contained at least seven items. The arrangement used by him which most nearly corresponded to that used in this experiment is the one with a linear model and with the copy displaced to one side. Piaget placed his children who succeeded in copying the bead order in this arrangement in Stage IIB [23] which ranges from about 5½ to 7 yr. The results reported here are consistent with Piaget's empirical findings, since the lowest age group which showed a large measure of success in matching the order of linear sequences of seven items was aged 5-9 to 6-5.

While confirming Piaget's empirical findings in this particular, the results also show (Table 3) that sequences of seven items are substantially more difficult to match than shorter sequences. Those *S*s as young as 4-3 to 4-11 showed a considerable degree of success in ordering three items. It is possible that the greater difficulty of the sequence of seven was due, not to any lack of understanding of order per se, but to the greater demands made by the longer sequence upon the *S*s' attention, interest, and persistence, or to other factors of a similar nature. This is consistent with the *S*s' behavior with the sequence of seven. The older *S*s took longer to compare the sample with the choices available and were more thorough in their comparisons, often carefully matching the colors one by one and indicating them on the sample with their fingers.

Piaget's interpretation of his results is that a concept of order does not mature fully in children until around age 7 yr. The results of this experiment suggest that whatever logical ability is involved in these tasks, it is already present in many children before age five, and the possibility of a yet earlier development is not excluded.

The present experiment differed from Piaget's, not only in that the number of items in linear sequence was varied, but also in the use of relatively nonverbal tech-

[22] Suggested by Brewster Smith.

[23] Stage IIB is also defined by success with the circular-to-linear translation. The present study, which investigates only linear-to-linear transformations, neither corroborates nor disproves Piaget's views of the development of nonlinear order concepts. It has already been pointed out that Piaget gives no direct evidence which demonstrates that the same psychological process is involved in linear and nonlinear transformations.

niques. Where the sequence contained seven items, the fact that Piaget's results were comparable to those reported here, suggests that for these sequences it made little difference whether Piaget's verbal instructions or the relatively nonverbal matching-from-sample technique was used. This equivalence may well have been because the *S*s who solved the sequences of seven items were the older children who were probably better acquainted with the words used in the instructions. Since Piaget did not use sequences of less than seven items, it is unclear to what extent the nonverbal technique facilitated the solutions of the younger *S*s on these sequences.

The important effect of the use of the nonverbal method is that lack of language comprehension cannot be held responsible for the total failure of some *S*s to discriminate order on any of the series.

The criticism of Piaget which is suggested by the data is similar to the criticism made in the preceding chapter of his views on the development of the formal reasoning inherent in measurement. In his writings Piaget shows little sign of appreciating the fact that the age at which the responses he studies can be elicited is profoundly dependent on the specific experimental conditions he employs. When the ease with which a response can be elicited is greatly altered by varying certain task parameters, is it legitimate or profitable to attempt to state an age at which the response develops? In the case of order discrimination, might one not conclude that it is unsatisfactory to postulate a logical reasoning process which defines an "emergent level" of development, since one always has to consider how many objects are involved, perhaps also the nature of the objects, and how the task is presented?

While this line of attack is certainly not fatal for the notion of levels of development, it does demonstrate the necessity for systematic exploration of the factors that inhibit or facilitate solution of order-discrimination tasks. Obviously, if one seeks to state an age at which an order-discrimination response develops, the only age which is not completely arbitrary is the earliest age at which order discrimination can be unequivocally elicited using the simplest experimental procedures. One cannot know the simplest procedures until one has systematically explored all parameters which are not involved in the definition of order discrimination. The experiment reported is a contribution in this direction.

Systematic parametric studies will lead to one of the two following results:

First, it may be found that a point is reached at which no further simplification of procedure is possible, without so profoundly altering the nature of the task that "order" ceases to be the relevant cue. This point will presumably be reached when the extraneous factors whose intervention prevents the youngest *S*s from solving the task have been discovered and eliminated.

Second, it may be found that the concept of order discrimination disintegrates on analysis. That is, it may prove impossible to define "sequence" as a cue in terms of clear-cut experimental operations. There is nothing in the results to suggest that this will be the case and Piaget's concept of levels of development is therefore broadly confirmed. However, a conclusion of this nature should be advanced with caution, since there are many instances in the history of comparative psychology where distinctions between "levels" have been found untenable when analyzed.[24]

[24] An instructive example is the distinction once made by Bingham (1914, 1922) and Hunter (1913) between perception of "form, per se" and perception of "shape" or "pattern." When trained to discriminate two forms, "lower" organisms were believed to have learned to respond to a particular retinal distribution of light ("pattern" vision). "Higher" organisms, who freely generalized whatever they had learned to many stimuli possessing similar geometric properties to those on which they were trained, were said to be responding to "form per se." However, analysis demonstrated the unreality of this distinction. It was found that "generalization after learned discrimination between two stimuli is an obvious and universal phenomenon" (Andrew & Harlow, 1942, p. 18), and that response is never simply to a particular retinal distribution of light. Lashley concluded that there is "no evidence that pattern or shape are ever the basis of reaction . . . neither phylogenetic studies nor evidence from pathology reveals pattern vision as a primitive function" (Lashley, 1938, p. 138).

It is clear that parallel issues are raised by the results of the experiments described in this and in the preceding section. That is, of course, no accident: it is likely that the same issues will present themselves whenever any of the formal reasoning processes studied by Piaget are the subject of experiment. The theoretical question is whether the kinds of reasoning whose formal structure is described in logic and mathematics constitute emergent psychological processes or whether, when the experimental procedures used to elicit these kinds of reasoning are systematically varied, it is found that the distinctions between the "emergent" and other more "primitive" processes are not translatable into experimental terms so that clearly definable levels emerge.

Recapitulation

A study by Piaget which explores the development of children's concepts of position order was discussed. It was pointed out that order discrimination is probably a function of several variables unexplored by Piaget—for example, the number and nature of the objects in sequence, and vocabulary development when the test instructions employ the word "order," or synonyms.

An experiment was designed which employed the matching-from-sample procedure developed in work with animals, so that the results are unlikely to be affected by vocabulary development. The number of objects in the sequences to be matched was varied systematically from two to seven, so that the development of order-discrimination responses in children could be explored as a function of this variable.

It was found that with those sequences which contained approximately the same number of objects as the sequences used by Piaget (i.e., seven objects), results comparable to his were obtained. However, it was also found that when the number of objects in sequence was reduced, order discrimination was elicited at an age considerably earlier than the age at which Piaget claims it develops in children. The importance of the variable, number of objects in sequence, was therefore demonstrated. It is likely, though not proven, that correct responses among the younger *S*s were facilitated by the relatively nonverbal presentation of the task.

Piaget's concept of "levels" of development was discussed. The evidence indicates that in order to establish discrimination of sequence as an "emergent" process, the effects of varying a number of factors other than the "order" cue itself require exploration. However, there is nothing in the results to suggest that the relational cue of spatial sequence cannot be given clear experimental definition and the data support the view that response to this cue may constitute an "emergent." To this important extent Piaget's concept of "levels" is confirmed.

The Relation of Order Discrimination to the Inferential Response $A > B.B > C. \supset .A > C$

In the previous sections it was shown that, although Piaget may be mistaken about the specific ages at which order discrimination and the inferential response $A > B.B > C. \supset .A > C$ develop, there is nothing to indicate that his concept of levels of development should be rejected.

In this section and the following one, Piaget's theory of intelligence, which contains his definition of general levels of intellectual functioning, is discussed. This theory suggests specific predictions about the relationship between discrimination of order and the inferential response studied.

Piaget's Theory of Intelligence

According to Piaget, logical reasoning consists of a group of "emergent" intellectual functions. He refers to these as "operational thinking," with which he identifies "intelligence." His theory of intelligence, therefore, attempts to give a detailed definition of logical reasoning in the following manner.

Piaget's basic postulate is that mature intellectual processes are isomorphic with

the "operations" of bivalent logic.[25] An example may clarify this use of the term "isomorphic." "Addition" of numbers is an intellectual process occurring in children when they do arithmetic; it is presumed to be isomorphic with the "additive operation" discussed by mathematicians. Similarly, deduction is an intellectual process (or processes) which can be placed in correspondence with one or more logical definitions. Piaget's definitions of the logical operations with which intellectual processes are assumed to be isomorphic are similar, but not identical, to the operations of the spatial, class, and propositional interpretations of the Boolean calculus, as described in texts on logic. Thus, the principal processes for which he finds isomorphic logical definitions are conjunction and disjunction of classes, of regions of space and of propositions,[26] and the corresponding relations of class inclusion, geographical or spatial inclusion, and implication. These processes are assumed to be, so to speak, the "atoms" of which intelligence is made. This part of Piaget's theory is essentially an essay in formal logic rather than psychology. According to Piaget, the relation between the logical definitions and the corresponding intellectual processes is analogous to the relation between pure geometry and physical geometry, or to the isomorphism between complex algebra and certain properties of alternating current electric circuits.[27]

Psychologically, the terms *conjunction* and *disjunction* are ambiguous. To illustrate the meaning intended here, class conjunction will be discussed in some detail. A conjunction of classes is that class whose members share the attributes of two other classes: thus, the class of white sheep is the conjunction of the class of white objects and the class of sheep, the class of balls (with exceptions) is the conjunction of the class of spheres and the class of playthings.

It might be argued that whenever we identify an object we perceive (e.g., a white sheep) we are performing a conjunction of classes, since any perceptual identification requires the simultaneous recognition of a number of different cues. If class conjunction is interpreted psychologically as the recognition of combinations of cues, it is obvious that it is an extremely primitive process which occurs early in phylogenetic development: many nonverbalizing organisms certainly discriminate classes of objects on the basis of combinations of cues. The apparent ubiquity of logical operations, viewed from this standpoint, appears to have led Lashley (1956) to assert that the operations of bivalent logic are coincident with the existence of nervous tissue, and this has led him to insist upon the need for a "neurology of logic."

Whether or not one selects a logical model for perception appears to be a matter of taste; many workers prefer a statistical model or one taken from information theory (e.g., Attneave, 1954; Binder, 1955). It is clear that whatever the reasoning processes are to which Piaget refers in his studies, they are *not* the processes involved in perceptual identification as discussed in the preceding paragraph. Unfortunately, the writer has not been able to find an adequate discussion by Piaget of this ambiguity in his concept of logical operation;

[25] It is difficult to phrase this postulate so as to avoid becoming embroiled in epistomological argument. The epistomology of the connection between reasoning process and logical operation is still disputed among logicians. To identify logical operation and reasoning process is Boole's fallacy and Piaget is well aware of this. The term "isomorphic" has been selected to express the postulated relation between logical operation and intellectual process in order—hopefully—to circumvent the epistomological issues. The use of this term does not do violence to Piaget's generally expressed views (Piaget, 1949, pp. 1–31; Piaget, 1950, pp. 18–37; Piaget, 1953b).

[26] The logical definitions of these operations may be found in most introductory texts in symbolic logic (e.g., Langer, 1953). Piaget's own definitions are expressed somewhat differently, but a review of these differences would complicate this discussion unnecessarily. See Piaget (1949; or 1950, pp. 43–47) for a brief summary of Piaget's system. For examples of these operations in ordinary speech, see Footnote 3.

[27] In a book review, Isaacs (1951) discusses this point in relation to Piaget's theory.

consequently, the interpretation of his conception that follows is the writer's.

In the English language, a conjunction of classes is denoted by a syntactical device: the qualification of a noun by an adjective as, for example, in the phrases "white sheep," "navigable rivers," etc. To render the following discussion more concrete, suppose that the English language be modified so that a conjunction of classes is expressed by the nonsense syllable "ZIK." In this modified English, white sheep would be denoted by the phrase "ZIK-whitethings-sheep"; navigable rivers would become "ZIK-navigablewaters-rivers"; balls would be denoted by the same word, "balls," but might also be represented by the phrase "ZIK-spheres-playthings." It is clear that the words "balls" and "ZIK" have quite different meanings. Knowledge of the meaning of the words "ball" and "white sheep," reflected in correct identification of these objects, is not the same as knowledge of the meaning of the word "ZIK," reflected in the proper usage of this term. The latter, let us say, reflects a general concept of class conjunction, which is not indicated by perceptual discrimination of either balls or white sheep. I believe it is to this, the discrimination of the meaning of the word "ZIK," that Piaget's theory refers when he uses the term *multiplication* or *conjunction of classes.*

There are many words and phrases in the English language which have a meaning as abstract as that of "ZIK" in the modified English discussed above. "Like," "similar to," "same as," "analogous to," "type of," "kind of," "species," "genus," "both . . . and," "either . . . or," "is a," "common to," are a few examples. These words express general characteristics of classes and have no reference to any specific class. The intellectual processes which Piaget's theory, as it is construed here, treats as isomorphic to the logical definitions of conjunction and disjunction of classes and of the relation of class inclusion are the kinds of processes which are reflected in the understanding and proper usage of such words and phrases.

The distinction made between the learning of names of objects and the learning of words such as "ZIK" is not new to psychology. It seems closely related to the distinction Lashley (1938) has made between "first-order generalizations" and "second-order generalizations." Examples in the literature of first-order generalizations are learning to discriminate "triangularity per se" (Andrew & Harlow, 1942; Fields: 1932, 1936; Gellerman, 1933; Lashley, 1938),[28] and to distinguish the identical elements in Hull's (1920) Chinese ideographs. Lashley's example of a problem whose solution requires second-order generalization is the "oddity" problem (Harlow, 1951; Lashley, 1938). The writer would suggest that other examples among nonverbal tests are the "sameness-difference" problem administered to chimpanzees (Robinson, 1955), and the experiments reported in the present paper. Examples of verbally communicated problems involving second-order generalization are provided in studies of the learning of logical relations by children (Kreezer & Dallenbach, 1929; Schooley & Hartmann, 1937).

It may be noted that the distinction between "abstract" and "concrete" thinking made by Goldstein and Scheerer (1941) appears to have much in common with Piaget's distinction between "operational" and "pre-operational" thinking and with Lashley's between first- and second-order generalizations.[29] All these distinctions appear to follow that made by logicians between "class" and "class of classes."[30]

Assuming the validity of the postulate that mature intellectual processes are iso-

[28] That is, learning to choose triangular stimuli, as against other shapes, irrespective of the size, color, or orientation of the stimuli.

[29] However, as in his discussion of amnesic aphasia, Goldstein (1948) sometimes appears to suggest that it is the nominal aspects of speech which are most clearly related to "abstract" thinking, whereas in Piaget's theory, it would be, of course, the syntactical aspects.

[30] Cf. Russell's (1920, Ch. 2) definition of number.

morphic with the operations of bivalent logic,[31] Piaget adopts the following research method. He sets children problems whose solution requires a certain line of logical reasoning. He then observes how children at various ages solve these problems, and he tries to observe the method of solution. Consideration of the ages of the children who solve the problem leads Piaget to specify the age at which the reasoning process involved in the problem develops.

He finds that certain types of problems are solved at the same age. Although he never explores the intercorrelations between success and failure on different problems, he concludes that reasoning processes develop in groups at particular ages. Specifically, he claims that the processes of class conjunction, class disjunction, and the concept of the relation of class inclusion develop in children at approximately 7 yr. of age. He also claims that the analogous operations of the spatial interpretation of Boole's system[32] develop at the same age. At around 11 yr., the child becomes capable of reasoning with propositions. That is, the processes of conjunction and disjunction of propositions develop at this age along with the concept of the relation of implication.

It can be seen, therefore, that Piaget's notion that reasoning processes develop in groups clearly implies the postulate that where operations, inferences, etc., are mutually interdependent (i.e., the logical definition of the one involves reference to the other, or that both derive from the same interpretation of a logical calculus), then the corresponding reasoning processes develop in association in children's thinking.[33] Thus, the operations of class conjunction, class disjunction, and the concept of the relation of class inclusion all derive from the class interpretation of the Boolean calculus and would therefore be expected to develop in children as a group. Later it will be shown that position order and the proposition $A > B.B > C. \supset .A > C$ (where ">" is interpreted "longer") are logically interdependent and, as an illustration of how Piaget's theory might operate, this postulate will be used to predict a high correlation between success in the tasks described in the two previous sections.

Summarizing Piaget's theory schematically, it consists, first, of a set of logical definitions with which the intellectual processes subsumed under the title "intelligence" are postulated to be isomorphic. Second, it postulates implicitly that where the logical definitions are interdependent or derive from the same interpretation of a logical system, the corresponding intellectual processes will be found to develop in association in children's thinking. Third, the theory contains a group of empirical laws stating the ages at which certain groups of processes develop. One may note that the ages indicated in these empirical laws could readily be altered where necessary without changing the main structure of Piaget's theory.

A criticism of Piaget's concept of intellectual process. In theory construction in psychology, intellectual processes have the theoretical status of hypothetical constructs. A hypothetical construct requires two kinds of specification. First, it is specified in terms of its relations to other constructs of a theoretical system. In Piaget's theory this kind of definition is contained in the two postulates that intellectual processes are isomorphic to definitions of logical op-

[31] One may note that there is no a priori reason why bivalent logic, rather than any of the polyvalent logics, should be selected as the model for reasoning.

[32] As noted in Footnote 3, Piaget's system differs somewhat from Boole's, but the differences can be neglected for the purpose of this study.

[33] The proposition that logical and ontogenetic priority correspond would seem to be a corollary. That is, where one logical definition is dependent on another, the intellectual process corresponding to the latter develops earlier. That there is at least some truth in this proposition is suggested by many educational practices based on common sense: for example, children are customarily taught to add numbers before they are taught to multiply, presumably because multiplication logically presupposes addition.

erations, and that when the logical definitions are interdependent the corresponding intellectual processes develop in association. However, in modern behavioral theory a hypothetical construct of this kind requires specification of the experimental operations through which it may be elicited, and of the characteristics of the response which testifies to its presence; without this it lacks empirical reference. Unfortunately, there is little in Piaget's theory which fulfills this function. The goal of his experiments is not to provide adequate specifiying procedures, but only to provide situations in which he can diagnose the character of intellectual processes through a series of acute observations of the *S*s' behavior and language. The tasks he uses are therefore not intended to serve as specifying experimental procedures, and the results from the experiments reported in the previous sections demonstrate that the tasks used by him in his studies of length measurement and of concepts of position order are inadequate for this purpose. It may be that Piaget's view that psychological processes do not need to be specified in terms of experimental procedures leads him to neglect the need for careful experimental design, with the result that he sometimes overlooks important variables.

This criticism exposes an important weakness of Piaget's theory—namely, that its roots in empirical fact are not entirely adequate. The power of his theory lies in the fact that it is potentially capable of giving a detailed and precise account of the development of intellectual functioning, and that numerous predictions may be derived from a few simple postulates.

Derivation of the Genetic Association of Order Discrimination With the Inferential Response A > B.B > C. ⊃ .A > C

As an example of how Piaget's theory might operate, this relationship will be derived using the postulate that when definition of logical operations are interdependent, the intellectual processes isomorphic to these operations develop in association. It follows from this postulate that, if there is a clear relation between the definition in logic of serial order and the proposition A > B.B > C. ⊃ .A > C, then discrimination of position order and the inferential response A > B.B > C. ⊃ .A > C develop together in children.

The relation in logic between "order" and rules of the form A > B.B > C. ⊃ .A > C is clearly stated by Russell in a discussion of the definition of order. Russell (1920) states, "We must not look for the definition of order in the nature of the set of terms to be ordered. . . . The order lies, not in the *class* of terms, but in a relation among the members of the class, in respect of which some appear as earlier and some as later" (pp. 30–31).

According to Russell, ordering relations are asymmetric, transitive, and connected.[34] That is, the property of transitivity of a relation is part of the definition of order, or, as Nagel (1930) states, it is an "axiom" of order. The rule, A > B.B > C. ⊃ .A > C merely states the transitive nature of the relation "longer than."

Given that a relation is asymmetric,[35] and disregarding the property of connectedness,[36] one may therefore predict that as

[34] A relation is asymmetric if, holding between x and y, it does not also hold between y and x (e.g., longer than, left of, father of). A relation is transitive if, holding between x and y, and also between y and z, it holds between x and z (e.g., longer than, left of, same as). A relation is connected if it has the property that, "given any two terms of the class to be ordered, there [is] one which precedes and another which follows. For example, of any two integers, . . . one is smaller and the other greater; but of any two complex numbers this is not true" (Russell, 1920, p. 32).

[35] I have been unable to think of any experimental operations for investigating an *S*'s understanding of the asymmetry of a relation, which differ from operations used to train him to discriminate the relation.

[36] Connectedness is a property of the number series in which Russell was principally interested. It is not, however, a property of all series: for example, it is a property of positions on a scale (no two positions being identical), but it is not necessarily a property of objects measured against a scale (since two of these may be the same). In the experiment reported, the *S*s were presumably comparing lengths of objects, not positions on the length scale.

the relation becomes transitive for a child, he should begin to distinguish different sequential arrangements of objects ordered with respect to it.

A particular kind of series is specified by its generating relation. Thus, the relation "heavier than" gives rise to a series of weights. The relations "after" or "before" give rise to temporal series. The relation of class inclusion gives rise to a class hierarchy (e.g., Greeks: men: mortal beings: etc.). Similarly, in a position sequence, such as that involved in the order-discrimination task, the relevant relation is spatial: "left of," or "right of." The asymmetric transitive relation involved in the length-measurement task is also a spatial relation, but a different one, namely, "longer than" (or "shorter than"). In order, therefore, to predict that the development of a grasp of the transitivity of the relation "longer than" will be associated with the genesis of discrimination of position sequences ordered with respect to the relation "left of" or "right of," it is first necessary to show that the relation "longer than" can be defined in terms of the relations "left of" and "right of." However, the relation "longer than" can indeed be so defined: the statement "OA is longer than OB" is a verbal shorthand for some more awkward statement such as, "When OA and OB are juxtaposed so that their left hand extremities, O, are coincident, then the relative positions of their other extremities will be such that OA overlaps OB, i.e., such that A is to the right of B." Thus, relative length is defined in terms of overlapping positions, and it is therefore possible to treat the relation "longer than" as a special case of the left-right relation.

It can therefore be predicted that discrimination of position order and the inferential response $A > B.B > C. \supset .A > C$ will be found to develop in association in children, or—to say the same thing in different words—the age at which the relation "longer than" becomes transitive for children should be closely related to the age at which different serial arrangements of a set of objects are distinguished.

A machine model. The derivation of this prediction might be represented more concretely in the following manner. Suppose we construct a machine which will perform the inference $A > B.B > C. \supset .A > C$. Into one end of this machine we feed the relations between A and B, and between B and C; the machine performs the inference and we read off its deductions concerning the relations between A and C. Now let us construct a second machine which will perform the order discriminations of the order task. Russell's logical analysis of order informs us that there is no operation performed by the first machine which the second machine cannot also perform. Moreover, the first machine can perform certain operations of the second machine: it can discriminate the sequence of three objects.

In the two previous sections, some similarities between the behavior of children and of these hypothetical machines were investigated as a function of age. Logical analysis demonstrates that these two machines are in important respects two versions of the same machine. Consequently, when an *S* behaves like one of these machines, he should also behave like the other. One may therefore make the two following predictions: First, the ages of first appearance of order discrimination and of the inferential response $A > B.B > C. \supset .A > C$ will be found to be closely related. Since the age of first appearance is difficult to establish more than approximately, the second prediction is more significant: there should be an extremely high correlation between the two tasks.

Empirical Evidence

Figure 7 portrays the increase with age of the percentage of *S*s who made an inferential response either in Phase 2 or Phase 3 (i.e., without or with the supplementary verbal instructions) of the measurement task. Juxtaposed with this is the graph showing the increase with age of the percentage of *S*s who reached the criterion in any of the series of the order task. The latter curve is identical to that for Series C with three colors per card, since all *S*s

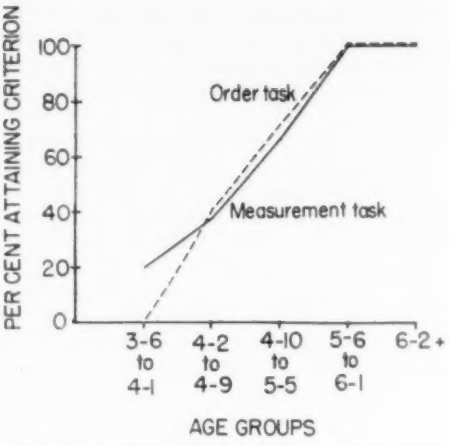

FIG. 7. Percentage of *S*s making an inferential response in the measurement task and discriminating order. The number of *S*s in each age group was 5, 13, 9, 8, and 5, respectively, for the measurement task, and 5, 12, 11, 11, and 5, respectively, for the order-discrimination task.

who reached the criterion in any series were successful in Series "C."

It can be seen that the course of development is almost exactly the same on the two tasks. The youngest age group is the only one at which there is a difference of more than a few percentage points between the two curves and this difference is caused entirely by one *S* who made an inferential response in the measurement task but failed to show any concept of order.[37]

[37] The age of this child was 3-8 and her adjustment to the order task was so unusual that the task cannot be regarded as a fair test of order discrimination for her. After the measurement task she discovered that other *S*s, who had not eaten the candy they had won, had had their earnings wrapped up for them in a piece of paper. Every piece of candy she won she then wanted wrapped up at once in a similar piece of paper and placed some distance away from her so that she would not be tempted to eat it immediately. Her attention was focused much more on what was going to happen to the candy than on the task. Once she had won three or four pieces responding randomly she was content and wanted to leave the conflictful experimental situation. Little significance can be attributed to her failure to reveal any concept of order.

The fact that the two developmental curves shown in Fig. 7 are almost superimposable confirms the prediction, but it would be dangerous to attach too much significance to it. Future investigation may well demonstrate that with other experimental procedures, either or both of the responses studied can be elicited in children of average intellectual capacity younger than these *S*s. If such should be the case, the close correspondence between the two curves would be revealed as fortuitous.

More significant is the fact that there was a strong tendency for *S*s either to fail both tasks or to pass both. This tendency occurred in all age groups. Altogether 37 *S*s took both tasks. Of these, 12 failed both, 22 showed some measure of success in both, 2 made an inferential response in the measurement task but completely failed to discriminate order, and one failed to make an inferential response but reached the criterion in some of the series of the order task. Even when one eliminates the two oldest age groups, in which 100% of *S*s passed both tasks, and the youngest age group, in which almost all *S*s failed both tasks, the correlation remains high. Considering, then, only the children aged 4–2 to 5–5, it was found that of the 20 children who took both tasks, 9 failed both, 9 were successful in both, 1 made an inferential response but showed no concept of order, and 1 showed some concept of order but failed to make an inferential response. Chi square = 12.8 ($p < .0001$), and the correlation (tetrachoric) is .96 ($\phi = .80$). In fact, in view of the impossibility of constructing tests of perfect reliability, the correlation appears to be as high as can be expected in careful psychological research.

Thus both predictions are substantially fulfilled. Discrimination of position order and the inferential response A > B.B > C. ⊃ .A > C were found to have the same course of development with age, at least in this sample of children and under the experimental conditions employed. Also, children who made an inferential response in the measurement task were also found to discriminate order, and vice versa.

Discussion

Colleagues with whom these results have been discussed have sometimes made the point that some degree of correlation between the tasks would be expected, since success on both tasks is probably correlated with intelligence test scores. In purely statistical terms, this point is no doubt valid. However, these tasks were designed in order to test a prediction derived from Piaget's theory of the development of intelligence and, in terms of Piaget's theory, success on them, by definition, reflects intelligence. It therefore begs the question to invoke intelligence to explain the correlation found, since the nature of intelligence is itself the subject of the investigation.

For this reason, intelligence tests were not administered to the *Ss*. One would, of course, expect that success on both tasks would correlate with mental age. The kind of inference demanded in the measurement task has been regarded since Aristotle as a manifestation of intelligence; stringing a chain of beads in a given order is a recurrent test item on the Binet.

It is known that correlations between individual test items at the preschool level are typically low relative to correlations between items at higher age levels (Goodenough, 1954, p. 466). Considering the age of the *Ss*, the correlation found is thus higher than one would expect on the general grounds that all tests of intellectual functioning correlate. However, the size of the correlation obtained is difficult to evaluate statistically because the sample is small and because it is not possible to establish confidence limits for a tetrachoric correlation or phi-coefficient.

One may note that Piaget's theory does not predict that all tests of intellectual functioning, paired indiscriminately, would correlate with each other. Between success in two carefully designed tasks such as the ones used here, one of which demands a grasp of the transitivity of the spatial relation "longer than," and the other of which demands the discrimination of sequences ordered with respect to an associated spatial relation, it predicts a very high correlation. But there appears to be nothing in Piaget's theory which would predict a particularly high correlation between success in these two spatial tasks and, for example, success in any task used to explore the development of operations of classification.[38, 39]

Conclusion

In the correlation found between order discrimination and the inferential response $A > B.B > C. \supset .A > C$, a prediction derived from a postulate implicit in Piaget's theory has been empirically verified. Within the restricted context of this study, the data confirm the validity of the postulate that when the definitions of logical operations are interdepedent, the isomorphic intellectual processes develop in association. Whether this postulate can be verified over a much wider range of children's intellectual responses and thus provide a solid empirical basis for Piaget's theory of the development of intelligence is a matter for further research.

Ages of Transition in the Development of Intelligence?

So far Piaget's theory has been examined only within the relatively limited context of the experimental work reported in the previous chapters.

---

[38] It is interesting to speculate that the distinction between the operations of the class and spatial interpretations of the logical calculus in Piaget's theory might provide a theoretical basis for the verbal and spatial factors usually found by factor analyses. It still appears to be a matter of interpretation in factor analysis whether the so-called verbal factor is truly a linguistic factor, or whether it relates to reasoning about class membership, relations between classes, etc., which provide the content of so many verbal tests.

[39] An unpublished study of the development of the class-inclusion relation, employed many of the *Ss* used in the experiments reported here. There was no obvious correlation between solution of the "class-inclusion" problem and success in the length-measurement and order-discrimination tasks. Unfortunately, the small number of *Ss* and the variety of factors found to be related to solution of the class-inclusion problem greatly limit the significance of this finding.

Piaget's theory affirms that in the course of development there are ages of transition during which intelligence changes from relatively "primitive" to more "mature" forms. Such important changes in the nature of intellect should be manifest in longitudinal studies of the development of intelligence as measured by intelligence tests. In the past twenty years there have been several studies of this kind, of which probably the best known is the Berkeley growth study (Jones & Bayley, 1941). In conclusion, the results from these studies are summarized and Piaget's theory is related to the evidence from them.

In the past there have been two principal viewpoints about the development of intelligence. The classical view is that there is no change in the nature of intelligence during development, that there are no qualitative differences in thought and reasoning between early childhood and later years, and that intellect develops through steady and continuous quantitative increments from birth to, at least, adolescence. This viewpoint found its clearest expression in the doctrine of the "constancy of the IQ." At the time of the heyday of the classical view about 25–30 yr. ago, the exponents of the opposite viewpoint were largely limited to Piaget's group at Geneva, although the psychoanalytic view has this in common with Piaget's: the theoretical dichotomy between the "primary processes" characteristic of infancy and the "secondary processes" which are thought to develop approximately at the time of resolution of the Oedipus complex necessarily implies a transition in the development of intellect.

Many of the early criticisms of Piaget's work were advanced before the classical viewpoint had been generally discredited. These criticisms, therefore, tended to assume that there were no changes in the nature of intelligence during development and consequently where Piaget saw a research problem, his critics saw either no problem or a pseudo-problem. Thus, not only were Piaget's early theories (Piaget: 1926, 1928, 1930a, 1930b) severely criticized, but also, unfortunately, the problems on which he was working were largely disregarded. Consequently, most of these critical studies (Bruce, 1941; Deutsche, 1937; Huang, 1943; Huang & Lee, 1945; Isaacs, 1930; Mead, 1932; Oakes, 1946)[40] have tended to be unfruitful; they have led neither to alternative solutions nor to reformulations of Piaget's research questions.[41] They have largely confined themselves to demonstrating, often convincingly, that Piaget's theories, as he stated them, had serious flaws and that the evidence adduced by him was often a good deal less than cogent. Piaget has himself admitted the validity of some of the criticisms of his early work (Piaget, 1953b, pp. 12–13).[42]

The following section summarizes the evidence from longitudinal studies for the existence of changes in the nature of intelligence during early childhood, which are not changes in quantity alone.

Longitudinal Studies of Intelligence

During the last twenty years there has been increasing evidence that scores on psychological tests administered during infancy have little or no predictive value for later intelligence. Table 4 summarizes, as a function of age at initial testing and of test-retest interval, the correlations obtained by 14 investigators who have followed groups of preschool children over a period of years into later childhood. These correlations are remarkably consistent. Tests administered prior to 2 yr. of age sometimes show low but significant correlations over short periods, but the relationships appear to vanish when the test-retest interval is increased.

It is not until around age 3 that significant (but usually low) correlations with scores in adolescence are obtained, and a high relationship with adolescent intelligence is not consistently found until the late preschool and first school years.

[40] There have also been generally confirmatory studies (Dennis, 1943; Dennis & Russell, 1940; Nagy, 1948; Russell, 1940). A number of studies of egocentric speech are reviewed by McCarthy (1954).

[41] With at least one important exception (Vigotsky, 1939).

[42] However, his admission is ambiguous: it seems to be retracted later in the same paragraph.

TABLE 4

SUMMARY OF TEST-RETEST CORRELATIONS OBTAINED IN LONGITUDINAL STUDIES OF INTELLECTUAL DEVELOPMENT

Investigator	*N*	Tests Used[a]	Test and Retest Age	*r*
Nelson & Richards (1938, 1939)	48	G × MP	6 mo. × 2 yr.	.37*
Bayley (1949, 1955)	27	C.1 × CPS	6 mo. × 3 yr.	.13
Cattell (1940)	49	C × SB	6 mo. × 3 yr.	.34*
Furfey & Muehlenbein (1932)	27	LH × SB	6 mo. × 4¾ yr.	−.11
Anderson (1939)	91	G et al. × SB	6 mo. × 5 yr.	−.06
Bayley	27	C.1 × WB	6 mo. × 18 yr.	−.24
Wittenborn (1956)	196	G × SB	3–14 mo. × 4–9 yr.	.05
Nelson & Richards	45	G × SB	1 yr. × 3 yr.	.33
Cattell	100	C × SB	1 yr. × 3 yr.	.56*
Bayley	27	C.1 × CPS	13 mo. × 3 yr.	.41
Furfey & Muehlenbein	27	LH × SB	1 yr. × 4¾ yr.	−.20
Anderson	91	G et al. × SB	1 yr. × 5 yr.	.05
Bayley	27	C.1 × WB	13 mo. × 18 yr.	−.14
Skodak & Skeels[b]	43	KB × SB	13–21 mo. × 5½–8 yr.	.31
Skodak & Skeels[b]	51	KB × SB	13–21 mo. × 11–16 yr.	.18
Honzik (1938); Honzik, MacFarlane, & Allen (1948)	200	CPS × CPS	21 mo. × 5 yr.	.39*
Honzik et al.	150	CPS × WB	21 mo. × 18 yr.	.07
Anderson	91	G et al. × SB	2 yr. × 5 yr.	.45*
Cunningham (1934)	20	KB × SB	2 yr. × 8 yr.	.24
Bayley	27	CPS × WB	2 yr. × 18 yr.	.39
Driscoll (1933)	39	KB × SB	2–3 yr. × 5–9 yr.	.66*
Goodenough & Maurer (1942)	26	MPS × SB	1½–3 yr. × 7½ yr.	.26
Goodenough & Maurer	21	MPS × SB	1½–3 yr. × 12½ yr.	.18
Bradway (1944, 1945)	52	SB × SB	2–3½ yr. × 13 yr.	.66*
Katz (1941)	268	SB × SB	3 yr. × 5 yr.	.62*
Ebert & Simmons (1943)	181	SB × SB	3 yr. × 6 yr.	.55*
Honzik et al.	200	CPS × SB	3 yr. × 6 yr.	.57*
Ebert & Simmons	181	SB × SB	3 yr. × 10 yr.	.50*
Honzik et al.	200	CPS × WB	3 yr. × 18 yr.	.35*
Bayley	27	CPS × WB	3 yr. × 18 yr.	.40
Skodak & Skeels[b]	34	KB × SB	2½–3½ yr. × 11–16 yr.	.51*
Goodenough & Maurer	28	MPS × SB	3–4 yr. × 12½ yr.	.61*
Ebert & Simmons	181	SB × SB	4 yr. × 10 yr.	.59*
Bayley	27	CPS × WB	4 yr. × 18 yr.	.52*
Bradway	86	SB × SB	4–5½ yr. × 15 yr.	.67*
Bayley	27	SB × WB	6 yr. × 18 yr.	.77*
Honzik et al.	200	SB × WB	7 yr. × 18 yr.	.71*

[a] Abbreviations for tests: G—Gesell Developmental Schedule; MP—Merrill-Palmer Performance Scale; C.1—California First-Year Mental Scale; CPS—California Preschool Scale; C—Cattell Infant Intelligence Scale; LH—Linfert-Hierholzer Infant Scale; MPS—Minnesota Preschool Scale; KB—Kuhlmann-Binet Scale; SB—Stanford-Binet Scale (1916 or 1937 revision); WB—Wechsler-Bellevue Intelligence Scale.

[b] Skodak and Skeels do not report correlations by age groups; the writer has computed these correlations from their published data (Skodak & Skeels, 1949, Table 17, pp. 122–24).

* Significant at the .01 level.

Several reasons have been proposed for the lack of correlation of early test scores with those of later ages. Low reliability cannot be a factor since after 6 mo. of age reliability is usually adequate (Cattell, 1940, p. 49). Escalona (1950) has suggested that the predictive power of infant tests could be improved by adjusting scores on the early tests according to clinical judgments, so that the scores represent not simply the sum of the items passed but clinical estimates of the child's "potential" at the time of testing. She presents some evidence that somewhat higher correlations might be obtained by this method over a short period of years. However, Wittenborn (1956) and his collaborators divided their *S*s into two groups. In one of these the infant intelligence classification was "unqualified"; in the other, the examiner expressed lack of confidence in the results. The correlations with later intelligence were about equally low in both groups.

Another explanation that has been advanced for the inconsistency of early test scores is that the infant tests do not contain suitable kinds of items. Catalano and McCarthy (1954) have suggested that measures of the prelinguistic vocalizations of infants might predict later intelligence better than the infant scale items. The early vocalizations of their 23 *S*s were sampled at an average age of 13 mo., and intelligence tests were administered after an average interval of 31 mo. The eight correlations reported varied from .04 (frequency of back vowels) to .45 (number of consonant types). They appear to be of about the same discouragingly small size as the correlations between comparable ages summarized in Table 4.

Several investigators (Bayley, 1955; Bradway, 1945; Maurer, 1946; Nelson & Richards: 1938, 1939) have performed item analyses of the infant scale test items. Bayley and her associates used the adolescent IQs of their 40 *S*s as the criterion variable. They selected the infant test items which most clearly differentiated the six brightest and six dullest children in adolescence. The total scores for the whole group of infants on these items were then correlated with the intelligence test scores of the same children at 14–16 yr. Bayley reports, "We were unable to get significant correlations even though our sample was composed in large part of the cases on whom the items were selected, including all of the extreme cases that would determine a relationship." She concludes, "I am inclined to think that the major reason for this failure rests in the nature of intelligence itself" (Bayley, 1955, p. 807).

This conclusion suggests strongly that different factors are involved in tests at different ages in early childhood and that there may be ages of transition during which the factors involved in later intelligence develop in children. Hofstaetter (1954) performed a factor analysis of the interage correlations of the Berkeley growth study and found the variance could be accounted for by three factors, with negligible residuals (see Table 5). To the first factor, Hofstaetter gives the name "sensori-motor alertness." He suggests the name "persistence" for the second factor, and relates it to the influence on the test results of such characteristics as cooperation and negativism. The third factor, Hofstaetter says "represents what is commonly called intelligence." Since these factors are derived from correlations between total scores on omnibus intelligence tests, no factors comparable to the spatial, verbal, etc., factors found in more conventional factor analyses can be expected.

Some caution must be exercised in interpreting these factors in terms of psychological functions. To equate one of them with an "ability" involves the dangerous assumption that there are no individual differences in the rate of increase of the ability over the period of high factor loading: otherwise, the test placement of the children with respect to each other would alter and the factor would not appear. It is not clear whether Hofstaetter conceives of "sensori-motor alertness" as an "ability."

TABLE 5

FACTOR LOADINGS (AFTER ROTATION) OF INTELLIGENCE TESTS AT VARIOUS AGES, REPRODUCED FROM HOFSTAETTER'S[a] ANALYSIS OF THE INTERAGE CORRELATIONS OF THE BERKELEY GROWTH STUDY

Factor	Test Age in Months												
	2	5	8	11	14	21	32	48	72	108	144	180	210
No. 1	45	77	88	87	80	53	35	03	00	02	08	00	08
No. 2	−48	−28	−13	15	46	54	65	57	36	07	02	00	05
No. 3	03	−04	06	18	10	39	50	68	90	97	96	98	94

[a] Hofstaetter, 1954, Table 2.

At this stage of knowledge, it seems wiser to refrain from interpreting these factors in terms of specific functions. They should probably be regarded in a more general manner as indicating that the kinds of processes which are involved in passing test items vary from age to age in a more or less regular manner during the preschool period. The precise nature of these processes can only be established by further research, which will not necessarily employ factor-analytic methods.

The third factor does not exert a preponderant influence until 4 yr. of age and the loading of this factor increases considerably between 4 and 6 yr. However, the *Ss* of the Berkeley study constitute an intellectually superior group. From 6 yr. of age onwards, successive measures of their mean IQ on different tests range from about 116 to 132.[43] One would expect the intellectual development of a group of children of more nearly normal intelligence at maturity to be slower; moreover, the third factor may be related to the kind of test item rather than to age, and an average child is examined on a given item at a later age than one who is superior. It is therefore quite possible that in a normal group the high loadings on the third factor might emerge at a somewhat later age than in the *Ss* of the Berkeley study.[44]

The evidence suggests that there are probably several periods of transition during which the nature of intellectual functions change. Whatever function or group of functions is measured by the infant tests becomes less and less important as age increases. Whether these functions actually decline in importance, or whether the content of the later tests is simply irrelevant to them, is an open question. If there is no decline of these functions then, at least theoretically, it is possible to construct tests at later ages which would correlate with the infancy tests.[45]

Piaget's Levels and the Evidence From Tests

When one attempts to assess Piaget's views in the light of the above material, one finds some things that are predictable from his theory and others that are not. Among the latter one notes that there appear to be periods of transition which he does not mention.

In two works, Piaget (1952, 1954) has discussed in considerable detail the cognitive development that takes place during the first two years of life, his data consisting of observations made on his own three children. His conception of the child's world at birth seems in many ways similar to James' famous concept of a "booming buzzing confusion." The main achievement of the sensori-motor development of the first two years appears to be the transformation of this confusion into a relatively stable world of concrete objects in which the child situates himself. Piaget apparently believes the child's perceptual organization of his immediate environment to be largely complete before 2 yr. of age. Temporally, this period corresponds fairly well with the period of influence of Hofstaetter's first factor.

There is then an hiatus in Piaget's studies of child development. In his works on infancy he leaves the child 18 mo. to 2 yr. old; in his studies of "operational" thinking the youngest *Ss* are about 4 yr. old. The intermediate period of development has not—at least not yet—been extensively studied by Piaget. This hiatus corresponds

[43] Estimated from Bayley (1955, Fig. 8, p. 812).

[44] As yet, no such intellectually normal group has been followed over a period of years. In all 14 investigations whose results are summarized in Table 4, the mean IQ of the *Ss* was above average, and usually substantially above average, although in most samples the dispersion of the IQs around the group mean was normal. The socioeconomic status and education of the adults with whom the children lived were also above average.

[45] It would be interesting to know if, having been unable to construct infant tests which will predict later intelligence, some of the investigators cited have tried to find out if there are any kinds of items in the later tests which are consistently associated with early precocity. It is conceivable that by this means some light might be thrown on the nature and fate of the processes involved in the infant tests.

fairly well to the period of maximum influence of Hofstaetter's second factor.

In Bayley's sample of children the third factor first becomes noticeable at 21 mo., becomes the most important factor at 4 yr., and achieves a fairly stable ceiling at 6 yr. The long period of emergence of this factor is thus completed by approximately the age at which Piaget claims that "operational" thinking emerges. Even making allowance for the superior intelligence of the Berkeley children, it is obvious that this factor plays a considerable role in intelligence tests at an earlier age than the 6–8 yr. at which Piaget places the intellectual changes he describes. Following Piaget's theory, one would predict that a factor analysis of longitudinal data would yield some factor similar to Hofstaetter's third factor. However, one would not expect this factor to become noticeable so early nor to rise quite so fast. Thus, if Piaget's view of the nature of intelligence is valid, the age levels he reports are probably placed too late. This conclusion is consistent with the findings of the experiments reported. The inferential and order-discrimination responses studied in the earlier sections developed in the sample of children employed (whose average IQ is probably 20–30 points lower than that of the Berkeley group) at approximately the age at which one might expect whatever is measured by Hofstaetter's third factor to exert an important influence on their intellectual performance.

This correspondence suggests that when some of the intellectual processes upon which Piaget's theory bears are studied by more carefully designed methods, the age of transition found is quite consistent with published facts about the development of intelligence.

The longitudinal studies provide no evidence to confirm the existence of the period of transition which Piaget places at about 11–12 yr. of age. No fourth factor appears in Hofstaetter's analysis, which indicates that the rank order of children on omnibus tests does not tend to alter at this age. However, this fact cannot be regarded as definitive proof that there is no qualitative change in intellectual performance at this time, since there may be qualitative changes which are not associated with changes in rank order.[46] Whether or not there are changes in the factorial structure of intelligence at this age manifested in more conventional factor analyses, is another question. One may note that there is little evidence for this in the literature: the evidence is only that the spatial, verbal, and numerical factors become more distinct with increasing age (Garrett, 1946). However, the test batteries used do not usually employ tests properly suited for testing Piaget's view about the transition alleged to take place at age 11 yr.; Piaget believes that it is not until this age that children begin to reason about propositions. It might be interesting to examine as a function of age the factorial structure of a battery which includes tests demanding inferences from propositions.[47] In some tests of this kind which have been administered, children under 11 yr. have shown little success (Moore, 1929).

It is clear from the longitudinal studies that Piaget's view that there is an alteration in the nature of intellectual processes toward the end of the preschool years is not as far-fetched as some of his early critics seem to have believed, although it is also clear that Piaget places too late the age at which the transition takes place. This change can be readily explained in terms of the concept of "levels" of intellectual development and is rather difficult to account for by any other theory. In fact, the longitudinal studies provide probably the strongest available evidence in favor of the type of approach which postulates and attempts to identify, "emergent" intellectual processes during early childhood.

The results from the longitudinal studies do not, of course, prove the validity of Piaget's view of the nature of the transition

[46] It has, of course, been assumed that the converse is true: when there is a marked change in children's test placement relative to each other, it is reasonable to assume that some important qualitative alteration in performance has occurred.

[47] Moore and Anderson (1954) present an interesting use of the "calculus of propositions" in the design of difficult tests involving inferences from propositions.

during the middle and late preschool period. But they do demonstrate that the study of the nature of intellectual processes at different ages is concerned with a real, not a pseudo-problem. The main criticism that can be leveled in retrospect against Piaget's early critics is that they contributed little to this problem.[48] It may be suggested that Piaget's theory of intelligence warrants much more detailed investigation than it has hitherto received.

Summary and Conclusions

Piaget's theory of the development of intellect has been examined in a number of ways. The first and primary purpose was to select two areas of intellectual functioning discussed by Piaget and to subject to detailed investigation his description of the development that takes place during childhood within these areas. This purpose was accomplished in the second and third sections. In the second section, the development in children of certain logical inferences, whose validity is tacitly assumed when length is measured, was studied. In the third section, the development of children's concepts of position order was investigated. In each case Piaget's theory, his research methods, and the nature of his evidence were reviewed and discussed, and an experiment was conducted to evaluate certain weaknesses in his approach, suggested by the discussion.

The second purpose was to discuss Piaget's theory of the development of intelligence, to derive an hypothesis from it and to test this hypothesis. The hypothesis concerned the expected correlation between a child's performance of one of the inferences basic in the logic of measurement, and his possession of a concept of position order, reflected in the discrimination of sequences of objects arranged in different orders.

The third purpose was to assess in a general manner the concordance of Piaget's concept of levels of development with the available data on the development of intelligence as measured by intelligence tests.

The evidence suggests the following conclusions.

1. The inference $A > B.B > C. \supset .A > C$ (where ">" is interpreted "longer than"), which has the status of an axiom in the logic of length measurement, is generally available to children at least two years before the age at which Piaget locates its development. Factors other than correct reasoning according to the axiomatic standards of the logic of measurement therefore appear to be involved in his studies. The difference between Piaget's experimental procedures and those used here suggests that these factors are the development of skill and interest in the techniques of length measurement, and perhaps vocabulary development (especially, comprehension of the meanings of words and phrases such as "measure," "same length," etc.). The effect of these factors is probably to conceal the reasoning ability of many of Piaget's *S*s.

2. Children show some concept of position order, in general, at least two years before the age at which Piaget places the development of order discrimination. It was found that the number of items in the sequences to be discriminated is an important variable disregarded by Piaget. It is likely, though not proven, that discrimination of order is facilitated for the younger *S*s and the shorter sequences (i.e., less than seven items) when the task is presented using relatively nonverbal methods.

3. There is nothing in the data which suggests that Piaget's concept of "emergent levels" of development should be rejected. The objection that the apparent "age of emergence" is a function of the particular

[48] Of course, Piaget's early work was on egocentricity, animism, etc., and not explicitly on intelligence. However, Piaget always explained these phenomena as manifestations of the nature of thought and reasoning at the age at which he found them. Obviously the development of a more mature form of intelligence (which, as defined by Piaget, seems to be closely related to what the psychoanalysts call "secondary processes," Goldstein calls "abstract attitude," and Lashley, "generalizations of the second order") is likely to be associated with many changes in modes of communication, in the perception of causality, etc. This must be expected whether or not Piaget's specific views on egocentric speech, animism, etc., are valid.

task manipulanda and discriminanda was considered in some detail. This objection is not fatal for a concept of levels although it suffices to show that the problem of establishing levels is somewhat more complicated than Piaget appears to have envisaged. The reason why Piaget's results differ from those obtained here is not, in the writer's opinion, because his notion of levels is inapplicable, but because in his research design he fails to eliminate important variables which are not involved in the definition of the processes he sets out to investigate, and which hinder many of his *S*s from responding to the appropriate cues.

4. Piaget's theory of the development of intelligence predicts accurately the high correlation in the individual child between the availability of the inference $A > B.B > C. \supset .A > C$ and the possession of a concept of position order.

5. The available data on the development of intelligence as measured by intelligence tests suggest that there are periods of transition during the preschool years during which the nature of intellectual functions changes. These data support the notion of phases or levels in the development of intellect and seem rather difficult to account for on any other basis. However, the period of transition most relevant to this examination of Piaget's theory appears to occur during the middle and late preschool years, rather than at approximately age 7 yr., as Piaget reports. The evidence thus suggests, once again, that Piaget probably places too late the development of the intellectual processes upon which his theory bears.

While discussing Piaget's theory of intelligence, it was noted that, although his theory is in some ways admirably precise and clear, his conception of psychological process is at variance with that of modern behavioral theory. For Piaget, intellectual processes are not hypothetical constructs, which require specification in terms of the experimental procedures by which they may be elicited, but realities which need only to be diagnosed through intuitive analysis of an *S*'s behavior and language. In his research his major effort is therefore expended on analysis of his *S*s' methods of solution of problems, rather than on experimental design. This method of approach perhaps leads him to overlook relevant variables, with the result that he is misled as to the age of development of the processes he is investigating. In this respect, one contribution of the present study is that it begins to provide more carefully designed experimental procedures for eliciting a few of the processes upon which his theory bears.

In view of the many criticisms that have been made of Piaget, it should be emphasized that the data reported provide considerable support for his notion of emergent levels and also, within the limited scope of these experiments, for his conception of the emergent processes as "operational" thinking.

REFERENCES

ANDERSON, L. D. The predictive value of infancy tests in relation to intelligence at five years. *Child Develpm.*, 1939, **10**, 203–212.

ANDREW, G., & HARLOW, H. F. Performance of Macaque monkeys on a test of generalized triangularity. *Comp. Psychol. Monogr.*, 1942, **19**, No. 3.

ARIETI, S. Special logic of schizophrenic and other types of autistic thought. *Psychiatry*, 1948, **11**, 325–338.

ATTNEAVE, F. Some informational aspects of visual perception. *Psychol. Rev.*, 1954, **61**, 183–193.

BAYLEY, NANCY. Consistency and variability in the growth of intelligence from birth to eighteen years. *J. genet. Psychol.*, 1949, **75**, 165–196.

BAYLEY, NANCY. On the growth of intelligence. *Amer. Psychologist*, 1955, **10**, 805–818.

BERGMANN, G., & SPENCE, K. W. The logic of psychophysical measurement. *Psychol. Rev.*, 1944, **51**, 1–24.

BINDER, A. A statistical model for the process of visual recognition. *Psychol. Rev.*, 1955, **62**, 119–129.

BINGHAM, H. C. A definition of form. *J. anim. Behav.*, 1914, **4**, 136–141.

BINGHAM, H. C. Visual perception of the chick. *Behav. Monogr.*, 1922, **4**, No. 4.

BOOLE, G. *An investigation of the laws of thought.* (1854 ed.) New York: Dover, 1953.

BRADWAY, K. P. IQ constancy in the revised Stanford-Binet from the preschool to the junior high school level. *J. genet. Psychol.*, 1944, **65**, 197–217.

BRADWAY, K. P. Predictive value of Stanford-Binet preschool items. *J. educ. Psychol.*, 1945, **36**, 1–16.

BRUCE, M. Animism vs. evolution of the concept "alive." *J. Psychol.*, 1941, **12**, 81–90.

CAMPBELL, N. R. *Physics: The elements.* Cambridge: Cambridge Univer. Press, 1920.

CATALANO, F., & MCCARTHY, DOROTHEA. Infant speech as a possible predictor of later intelligence. *J. Psychol.*, 1954, **38**, 203–209.

CATTELL, PSYCHE. *The measurement of intelligence of infants and young children.* New York: Psychological Corp., 1940.

CUNNINGHAM, B. V. Infant IQ ratings evaluated after an interval of seven years. *J. exp. Educ.*, 1934, **3**, 84–87.

DENNIS, W. Animism and related tendencies in Hopi children. *J. abnorm. soc. Psychol.*, 1943, **38**, 21–36.

DENNIS, W., & RUSSELL, R. W. Piaget's questions applied to Zuni children. *Child Develpm.*, 1940, **11**, 181–187.

DEUTSCHE, J. M. The development of children's concepts of causal relations. Univer. Minnesota, Inst. Child Welf. Monogr. Series. Minneapolis: Univer. Minnesota Press, 1937.

DRISCOLL, G. P. The developmental status of the preschool child as a prognosis of future development. *Child Develpm. Monogr.*, 1933, No. 13.

EBERT, E., & SIMMONS, K. The Brush Foundation study of child growth and development. I. Psychometric tests. *Monogr. Soc. Res. Child Develpm.*, 1943, **8**, No. 2.

ESCALONA, SIBYLLE. The use of infant tests for predictive purposes. *Bull. Menninger Clin.*, 1950, **14**, 117–128.

EVARTS, E. V., & NISSEN, H. W. Test of "the abstract attitude" in chimpanzees following ablation of prefrontal cortex. *AMA Arch. Neurol. Psychiat.*, 1953, **69**, 323–331.

FIELDS, P. E. Studies in concept formation. I. Development of the concept of triangularity by the rat. *Comp. Psychol. Monogr.*, 1932, **9**, No. 2.

FIELDS, P. E. Studies in concept formation. IV. A comparison of white rats and raccoons with respect to their visual discrimination of certain geometric figures. *J. comp. Psychol.*, 1936, **21**, 341–355.

FRAISSE, P., & VAUTREY, P. La perception de l'espace, de la vitesse, et du temps chez l'enfant de cinq ans. I. L'espace et la vitesse. *Enfance*, 1952, **5**, 1–20. (a)

FRAISSE, P., & VAUTREY, P. La perception de l'espace, de la vitesse, et du temps chez l'enfant de cinq ans. II. Le Temps. *Enfance*, 1952, **5**, 102–119. (b)

FURFEY, P. H., & MUEHLENBEIN, J. The validity of infant intelligence tests. *J. genet. Psychol.*, 1932, **40**, 219–223.

GARRETT, H. E. A developmental theory of intelligence. *Amer. Psychologist*, 1946, **1**, 372–378.

GELLERMAN, L. W. Form discrimination in chimpanzees and two-year-old children. I. Form (triangularity) per se. *J. genet. Psychol.*, 1933, **42**, 3–27.

GOLDSTEIN, K. *Language and language disturbances.* New York: Grune & Stratton, 1948.

GOLDSTEIN, K., & SCHEERER, M. Abstract and concrete behavior: An experimental study with special tests. *Psychol. Monogr.*, 1941, **53**, No. 2 (Whole No. 239).

GOODENOUGH, FLORENCE L. The measurement of mental growth in childhood. In L. Carmichael (Ed.), *Manual of child psychology.* (2nd ed.) New York: Wiley, 1954.

GOODENOUGH, FLORENCE L., & MAURER, K. M. *The mental growth of children from two to fourteen years: A study of the predictive value of the Minnesota Preschool Scales.* Minneapolis: Univer. Minnesota Press, 1942.

HARLOW, H. F. Primate learning. In C. P. Stone (Ed.), *Comparative psychology.* (3d ed.) New York: Prentice-Hall, 1951.

HILGARD, E. R. Methods and procedures in the study of learning. In S. S. Stevens (Ed.), *Handbook of experimental psychology.* New York: Wiley, 1951.

HOFSTAETTER, P. R. The changing composition of "intelligence": A study in T-technique. *J. genet. Psychol.*, 1954, **85**, 159–164.

HONZIK, M. P. The constancy of mental test performance during the preschool period. *J. genet. Psychol.*, 1938, **52**, 285–302.

HONZIK, M. P., MACFARLANE, JEAN W., & ALLEN, L. The stability of mental test performance between two and eighteen years. *J. exp. Educ.*, 1948, **17**, 309–324.

HUANG, I. Children's conception of physical causality: A critical summary. *J. genet. Psychol.*, 1943, **63**, 71–121.

HUANG, I., & LEE, H. W. Experimental analysis of child animism. *J. genet. Psychol.*, 1945, **66**, 69–74.

HULL, C. L. Quantitative aspects of the evolution of concepts. *Psychol. Monogr.*, 1920, **28**, No. 1 (Whole No. 123).

HUNTER, W. S. The question of form perception. *J. anim. Behav.*, 1913, **3**, 329–333.

ISAACS, N. Review of J. Piaget, *Traité de logique*. *Brit. J. Psychol.*, 1951, **42**, 185–188.

ISAACS, SUSAN. *Intellectual growth in young children*. London: Routledge & Kegan Paul, 1930.

JONES, H. E., & BAYLEY, NANCY. The Berkeley growth study. *Child Develpm.*, 1941, **12**, 167–173.

KATZ, E. The constancy of the Stanford-Binet IQ from three to five years. *J. Psychol.*, 1941, **12**, 159–181.

KREEZER, G., & DALLENBACH, K. M. Learning the relation of opposition. *Amer. J. Psychol.*, 1929, **41**, 432–441.

LANGER, SUSANNE K. *An introduction to symbolic logic*. (2nd ed.) New York: Dover, 1953.

LASHLEY, K. S. The mechanism of vision. XV. Preliminary studies of the rat's capacity for detail vision. *J. gen. Psychol.*, 1938, **18**, 123–193.

LASHLEY, K. S. Cerebral organization and behavior. Address to the Association for Research in Nervous and Mental Diseases, New York, 1956.

MCCARTHY, DOROTHEA. Language development in children. In L. Carmichael (Ed.), *Manual of child psychology*. (2nd ed.) New York: Wiley, 1954.

MAURER, K. M. *Intellectual status at maturity as a criterion for selecting items in preschool tests*. Minneapolis: Univer. Minnesota Press, 1946.

MEAD, MARGARET. An investigation of the thought of primitive children with special reference to animism. *J. Royal Anthrop. Inst.*, 1932, **62**, 173–190.

MOORE, O. K., & ANDERSON, S. B. Modern logic and tasks for experiments on problem solving behavior. *J. Psychol.*, 1954, **38**, 151–160.

MOORE, T. V. The reasoning ability of children in the first years of school life. *Stud. Psychol. Psychiat., Catholic Univer. of Amer.*, 1929, **2**, No. 2.

NAGEL, E. Measurement. *Erkenntnis*, 1930, **2**, 313–333.

NAGY, M. The child's theories concerning death. *J. genet. Psychol.*, 1948, **73**, 3-27.

NELSON, V. L., & RICHARDS, T. W. Studies in mental development: I. Performance on Gesell items at six months and its predictive value for performance on mental tests at two and three years. *J. genet. Psychol.*, 1938, **52**, 303–325.

NELSON, V. L., & RICHARDS, T. W. Studies in mental development: III. Performance of twelve-months-old children on the Gesell schedule, and its predictive value for mental status at two and three years. *J. genet. Psychol.*, 1939, **54**, 181–191.

OAKES, M. F. Children's explanations of natural phenomenon. *Teachers' Coll. Contr. Educ.*, 1946, No. 926.

PIAGET, J. *The language and thought of the child*. New York: Harcourt, Brace, 1926.

PIAGET, J. *Judgment and reasoning in the child*. New York: Harcourt, Brace, 1928.

PIAGET, J. *The child's conception of physical causality*. New York: Harcourt, Brace, 1930. (a)

PIAGET, J. *The child's conception of the world*. New York: Harcourt, Brace, 1930 (b)

PIAGET, J. *Les notions de mouvement et de vitesse chez l'enfant*. Paris: Presses Universitaires, 1946.

PIAGET, J. *Traité de logique*. Paris: Colin, 1949.

PIAGET, J. *The psychology of intelligence*. London: Routledge & Kegan Paul, 1950.

PIAGET, J. *The origins of intelligence in children*. New York: Int. Universities Press, 1952.

PIAGET, J. How children form mathematical concepts. *Scientific Amer.*, 1953, **189** (5), 74–79. (a)

PIAGET, J. *Logic and psychology*. Manchester: Manchester Univer. Press, 1953. (b)

PIAGET, J. *The construction of reality in the child*. New York: Basic Books, 1954.

PIAGET, J. *De la logique concrète de l'enfant à la logique propositionelle de l'adolescent*. Paris: Presses Universitaires, 1955.

PIAGET, J., & INHELDER, BARBEL. *La génèse de l'idée de hasard chez l'enfant*. Paris: Presses Universitaires, 1953.

PIAGET, J., & INHELDER, BARBEL. *The child's conception of space*. London: Routledge & Kegan Paul, 1956.

PIAGET, J., INHELDER, BARBEL, & SZEMINSKA, ALINA. *La géométrie spontanée de l'enfant*. Paris: Presses Universitaires, 1948.

PIAGET, J., & SZEMINSKA, ALINA. *The child's conception of number*. London: Routledge & Kegan Paul, 1952.

RIOPELLE, A. J., ALPER, R. G., STRONG, P. M., & ADES, H. W. Multiple discrimination and patterned string performance of normal and temporal-lobectomized monkeys. *J. comp. physiol. Psychol.*, 1953, **46**, 145–149.

ROBINSON, J. S. The sameness-difference discrimination problem in chimpanzee. *J. comp. physiol. Psychol.*, 1955, **48**, 195–197.

RUSSELL, B. *Introduction to mathematical philosophy.* (2nd ed.) London: Allen & Unwin, 1920.

RUSSELL, R. W. Studies of animism: II. The development of animism. *J. genet. Psychol.*, 1940, **56**, 353–366.

SCHOOLEY, M., & HARTMANN, G. W. The role of insight in the learning of logical relations. *Amer. J. Psychol.*, 1937, **49**, 287–292.

SKODAK, MARIE, & SKEELS, H. M. A final follow-up of one hundred adopted children. *J. genet. Psychol.*, 1949, **75**, 85–125.

SPEARMAN, C. *The nature of intelligence.* London: Macmillan, 1923.

TERMAN, L. M., & MERRILL, MAUD A. *Measuring intelligence.* Boston: Houghton Mifflin, 1937.

VIGOTSKY, L. S. Thought and speech. *Psychiatry*, 1939, **2**, 29–54.

VON DOMARUS, E. The specific laws of logic in schizophrenia. In J. S. Kasinin (Ed.), *Language and thought in schizophrenia.* Berkeley: Univer. California Press, 1944.

WITTENBORN, J. R. A study of adoptive children: I. Interviews as a source of scores for children and their homes. II. The predictive validity of the Yale Developmental Examination of infant behavior. III. Relationships between some aspects of development and some aspects of environment for adoptive children. *Psychol. Monogr.*, 1956, **70**, Nos. 1–3 (Whole Nos. 408–410).

(Accepted for publication May 29, 1958)